PATENT COOPERATION TREATY

From the INTERNATIONAL SEARCHING AUTHORITY

To:
HIMANSHU S. AMIN
AMIN & TUROCY, LLP
1900 E. 9TH STREET, 24TH FLOOR
NATIONAL CITY CENTER
CLEVELAND, OH 44114

PCT

NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL SEARCH REPORT AND THE WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY, OR THE DECLARATION

(PCT Rule 44.1)

	Date of mailing (*day/month/year*) 05 JUN 2007
Applicant's or agent's file reference MS306693.26	**FOR FURTHER ACTION** See paragraphs 1 and 4 below
International application No. PCT/US04/24139	International filing date (*day/month/year*) 27 July 2004 (27.07.2004)
Applicant MICROSOFT CORPORATION	

1. ☒ The applicant is hereby notified that the international search report and the written opinion of the International Searching Authority have been established and are transmitted herewith.

   Filing of amendments and statement under Article 19:
   The applicant is entitled, if he so wishes, to amend the claims of the international application (see Rule 46):

   When? The time limit for filing such amendments is normally two months from the date of transmittal of the international search report.

   Where? Directly to the International Bureau of WIPO, 34 chemin des Colombettes 1211 Geneva 20, Switzerland, Facsimile No.: (41-22) 338.82.70.

   For more detailed instructions, see the notes on the accompanying sheet.

2. ☐ The applicant is hereby notified that no international search report will be established and that the declaration under Article 17(2)(a) to that effect and the written opinion of the International Searching Authority are transmitted herewith.

3. ☐ **With regard to the protest** against payment of (an) additional fee(s) under Rule 40.2, the applicant is notified that:

   ☐ the protest together with the decision thereon has been transmitted to the International Bureau together with the applicant's request to forward the texts of both the protest and the decision thereon to the designated Offices.

   ☐ no decision has been made yet on the protest; the applicant will be notified as soon as a decision is made.

4. **Reminders**

   Shortly after the expiration of **18 months** from the priority date, the international application will be published by the International Bureau. If the applicant wishes to avoid or postpone publication, a notice of withdrawal of the international application, or of the priority claim, must reach the International Bureau as provided in Rules 90*bis*.1 and 90*bis*.3, respectively, before the completion of the technical preparations for international publication.

   The applicant may submit comments on an informal basis on the written opinion of the International Searching Authority to the International Bureau. The International Bureau will send a copy of such comments to all designated Offices unless an international preliminary examination report has been or is to be established. These comments would also be made available to the public but not before the expiration of 30 months from the priority date.

   Within **19 months** from the priority date, but only in respect of some designated Offices, a demand for international preliminary examination must be filed if the applicant wishes to postpone the entry into the national phase **until 30 months** from the priority date (in some Offices even later); otherwise, the applicant must, **within 20 months** from the priority date, perform the prescribed acts for entry into the national phase before those designated Offices.

   In respect of other designated Offices, the time limit of **30 months** (or later) will apply even if no demand is filed within 19 months.

   See the Annex to Form PCT/IB/301 and, for details about the applicable time limits, Office by Office, see the *PCT Applicant's Guide*, Volume II, National Chapters and the WIPO Internet site.

Name and mailing address of the ISA/ US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (571) 273-3201	Authorized officer Wei Zhen Telephone No. 571-272-2100

Form PCT/ISA/220 (January 2004) (*See notes on accompanying sheet*)

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference MS306693.26	**FOR FURTHER ACTION**	see Form PCT/ISA/220 as well as, where applicable, item 5 below.
International application No. PCT/US04/24139	International filing date (*day/month/year*) 27 July 2004 (27.07.2004)	(Earliest) Priority Date (*day/month/year*) 24 October 2003 (24.10.2003)
Applicant MICROSOFT CORPORATION		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 7 sheets.

☒ It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the Report**
   a. With regard to the **language**, the international search was carried out on the basis of:
      ☒ the international application in the language in which it was filed.
      ☐ a translation of the international application into _____, which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b))
   b. ☐ With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, see Box No. I.
2. ☐ **Certain claims were found unsearchable** (See Box No. II)
3. ☐ **Unity of invention is lacking** (See Box No. III)
4. With regard to the **title**,
   ☒ the text is approved as submitted by the applicant.
   ☐ the text has been established by this Authority to read as follows:

5. With regard to the **abstract**,
   ☐ the text is approved as submitted by the applicant.
   ☒ the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box No. IV. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. With regard to the **drawings**,
   a. the figure of the **drawings** to be published with the abstract is Figure No. 1
      ☒ as suggested by the applicant.
      ☐ as selected by this Authority, because the applicant failed to suggest a figure.
      ☐ as selected by this Authority, because this figure better characterizes the invention.
   b. ☐ none of the figures is to be published with the abstract.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US04/24139

Box IV TEXT OF THE ABSTRACT (Continuation of Item 5 of the first sheet)

The present systems and methods disclose a system for personalizing computer functionality. End-users are provided with tools to easily write rich and complex preferences, for example, by using a plurality simple IF-THEN propositional logic. The preferences are then transformed into queries and executed efficiently on structured data. Preferences that are satisfied then execute actions such as providing notification or storing data in a particular folder. Furthermore, according to an aspect of the invention, data, logic, events, inter alia, are all schematized, thereby enabling sharing of data between application components and across applications.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US04/24139

A. CLASSIFICATION OF SUBJECT MATTER

IPC: **G06F 9/45**(2006.01),**G06F 7/00**

USPC: 717/140,707/100

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 717/140,707/100

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

EAST, consulted with Todd Swaan for 101 issue

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

USPAT, US-PGPUB, IBM_TDB, EPO, JPO, DERWENT, USOCR

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	US Patent No. 6,745,180 B2 (YAMANOUE) June 1, 2004, column 1, lines 8-15; column 22, lines 5-8; column 6, lines 19-22; column 8, lines 49 into column 9, column 7, lines 34-40; Figures 1, 13, 16, and claim 38.[0008], [0009], [0011],[0206], [0254] and [0267].	1-5, 11-13, 18-21, and 23 ---------- 6,8,9,10,22
X	US 6,606,618 B2 (DELO) 12 August 2003,see column 2, lines 49-51; column 7, lines 15-28; and column 5, lines 38-40.	14-17

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

30 March 2007 (30.03.2007)

Date of mailing of the international search report

05 JUN 2007

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

Facsimile No. (571) 273-3201

Authorized officer

Wei Zhen

Telephone No. 571-272-2100

Form PCT/ISA/210 (second sheet) (April 2005)

PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

To:
HIMANSHU S. AMIN
AMIN & TUROCY, LLP
1900 E. 9TH STREET, 24TH FLOOR
NATIONAL CITY CENTER
CLEVELAND, OH 44114

PCT

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

(PCT Rule 43*bis*.1)

Date of mailing *(day/month/year)*	05 JUN 2007
Applicant's or agent's file reference MS306693.26	**FOR FURTHER ACTION** See paragraph 2 below

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US04/24139	27 July 2004 (27.07.2004)	24 October 2003 (24.10.2003)

International Patent Classification (IPC) or both national classification and IPC

IPC: **G06F 9/45**(2006.01)
USPC: 717/140

Applicant

MICROSOFT CORPORATION

1. This opinion contains indications relating to the following items:

- ☒ Box No. I — Basis of the opinion
- ☐ Box No. II — Priority
- ☐ Box No. III — Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- ☐ Box No. IV — Lack of unity of invention
- ☒ Box No. V — Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- ☐ Box No. VI — Certain documents cited
- ☐ Box No. VII — Certain defects in the international application
- ☐ Box No. VIII — Certain observations on the international application

2. **FURTHER ACTION**

If a demand for international preliminary examination is made, this opinion will be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1*bis(b)* that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220.

3. For further details, see notes to Form PCT/ISA/220.

Name and mailing address of the ISA/ US	Date of completion of this opinion	Authorized officer
Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (571) 273-3201	23 May 2007 (23.05.2007)	Wei Zhen Telephone No. 571-272-2100

Form PCT/ISA/237 (cover sheet) (April 2005)

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US04/24139

Box No. I Basis of this opinion

1. With regard to the **language**, this opinion has been established on the basis of:

☒ the international application in the language in which it was filed

☐ a translation of the international application into ____, which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

2. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application and necessary to the claimed invention, this opinion has been established on the basis of:

a. type of material

☐ a sequence listing

☐ table(s) related to the sequence listing

b. format of material

☐ on paper

☐ in electronic form

c. time of filing/furnishing

☐ contained in the international application as filed.

☐ filed together with the international application in electronic form.

☐ furnished subsequently to this Authority for the purposes of search.

3. ☐ In addition, in the case that more than one version or copy of a sequence listing and/or table(s) relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

4. Additional comments:

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US04/24139

Box No. V Reasoned statement under Rule 43 *bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Claims	6-10, 14-17, 21-22	YES
	Claims	1-5, 11-13, 18-20, 23	NO
Inventive step (IS)	Claims	NONE	YES
	Claims	1-23	NO
Industrial applicability (IA)	Claims	1-23	YES
	Claims	NONE	NO

2. Citations and explanations:

Please See Continuation Sheet

Supplemental Box
In case the space in any of the preceding boxes is not sufficient.

V. 2. Citations and Explanations:
Claims 1-5, 11-13, 18-20, and 23 lack of novelty under PCT Article 33(2) as being anticipated by US Patent No. 6,745,180, by Yamanoue.
As Per claim 1, Yamanouc discloses:
A preference evaluation system comprising: a data store component for storing schematized data;
Yamanouc's disclosure stores user preference data, see Yamanouc's column 1, Lines 8-15, "The present invention relates to a data supply controlling device, a data supplying method, a storage medium (a data store) storing a data supplying program, and a data supplying system, in which a system is used which searches for information (of books, for instance) using a data supplying device and provides search results to the user (user terminal) via the data supply controlling device in such a manner to provide information suitable for user's preference and interest by referring to user data". And see Yamanouc's column 29, lines 14-20, "Incidentally, Japanese Examined Patent Publication No. 2976219 discloses that user data is recorded, for instance, in a storage medium such as IC memory, and the user data stored in the storage media above is read by a terminal of the user when a search for commercial information is conducted, and then a host computer conducts an information search by picking up only required data from the user data." -- where the storage medium is as a the 'data store'; as to the schematized data, any "data, logic, events, inter alia, are all schematized", see the citation in the current application Abstract, and paragraph [0010], "Schematization is the structuring of data in well-known and well-defined patterns, which enables multiple applications to recognize and interact with each other."
a compiler to compile applications including end-user specified preferences and store them in the data store; and
See Yamanoue's column 22, lines 5-8, "After the data compiled or read in $64 or $65 is sent to the user terminal 1 by the data supplying means 41", and column 22,
lines 17-20, "if the user selects to seedata in the user terminal 1 (end-user specified preferences) in the manner of the data browse in $67, the data to be browsed is compiled ($68) and immediately transferred to the user terminal 1 ($69)"; it's inherent that data has to be compiled by a compiler.
an execution engine to evaluate preferences stored in the data store upon the occurrence of one or more events. See Yamanoue's Fig. 13, and column 6, lines 19-22, "In the arrangement, the process by means of the method of supplying data can be implemented by a computer (for example, a data supply controlling device) executing the program"; and column 28, lines 42-46, "the stored program may be arranged to be executed by an access of a microprocessor (not illustrated), or arranged so that the program is executed (execution engine to evaluate the stored preferences) by reading the stored program and then downloading the read program to a program storage of a delivery server and a receiving server.". Basically Yamanoue's disclosure as specified in his claim 38, "A storage medium (data store) for storing a data supply program executed (execution engine) by a computer to implement a method of supplying data, the method comprising the steps of: storing, in user data storage means, user data for each user who is to be supplied with information through a user

Supplemental Box
In case **the space in any of the preceding boxes is not sufficient.**

terminal (store data upon the occurrence of one or more events); generafing, based on at least the user data, a query for a data supplying device to search for the information; storing, in search result storage means, a result of a search conducted by the data supplying device for the information in accordance with the query".

As Per claim 2, Yamanoue discloses:
The system of claim 1,further comprising an action component for taking one or more actions specified by a conditionally valid preference.
For claim 1 feature see claim 1 rejection, conditionally valid preference is disclosed in Yamanoue's 'search rules', see Yamanoue's Fig. 1 and description, and example in column 8, lines 49 into column 9, "If the data server 2 is a server searching information of books as in the present embodiment, as the search rule described above, the search rule generator 24 sets a rule that enables the search service center 3 to generate queries, which are capable of searching information of books, on the basis of the user data.", couple 'IF' conditions are assessed (column 8, lines 56-62) from Data Server 2, further actions are taken at Search Service Center 3. The query rules specified in FIG. 3 can all considered as 'conditionally valid preference'.

As Per claim 3, Yamanoue discloses:
The system of claim 2, the action component comprising a notification component that transforms and formats notification data generated by the execution engine based on a user preference for one or more user communication devices.
For claim 2 feature see claim 2 rejection, Yamanoue's disclosure including a 'search result management' component, which would transform and form notification data generated by the execution engine based on a user preference on a user's communication device, see Yamanoue's FIG 1, and FIG. 16, and description in column 8, lines 12-21, "The search result management means 13 (a notification component) stores the search results being transferred from the search service center 3 to the user terminal 1" and "...The display 15 offers a GUI (Graphical User Interface) for various operations and displays search results, and includes a PC monitor, for instance."

As Per claim 4, Yamanoue discloses:
The system of claim 1, wherein the communication devices include a mobile phone, a pager, a PDA, and a computer.
For claim 1 feature see claim 1 rejection, Yamanoue's teaching include a computer or a mobile phone, see Yamanoue's column 7, lines 34-40, "The user terminal 1 is used by the user being provided data by the present system. What can be used as the terminal 1 are, for instance, devices owned by the users and can be connected to the network 4 such as a PC (personal computer), a mobile information terminal and a mobile phone, and also a dedicated terminal capable of being used in the system."

As Per claim 5, Yamanoue discloses:
The system of claim 1, further comprising an event component to extract event data from an event source and store the data in the data stork
For claim 1 feature see claim 1 rejection, each data storage or data query/retrieval in Yamanoue's disclosure is considered as an event, see column 3, lines 38-59, "A data supply controlling device in accordance with the present invention, in order to accomplish the foregoing objective, is characterized in that it includes: a user data storage section for storing user data for each user who is to be supplied with information through a user terminal" - the data supply controlling device serves as an event component, which can extract event data from the input source and store the data in the data store.

As Per claim 11, Yamanoue discloses:
The system of claim 1, the execution engine evaluates preferences by executing queries on data stored in the data stork
For claim 1 feature see claim 1 rejection, for rest of claim 11 feature see Yamanoue's Abstract, "A data supply controlling device comprises a data base for user data which stores user data matched with each user. The data base of user data can be queried (executing queries on data stored m the data store) in accordance with the user data so ~hat a data server performs a search according to the query and stores the search results in a search result data base." and column 4, lines 46-55, "storing. in a user data storage section, user data for each user who is to be supplied with information through a user terminal; generating, based on at least the user data, a query to search data supplied from a data supplying device; searching for information in accordance with the query; (executing queries) storing a search result in a search result storage section;"

As Per claim 12, Yamanoue discloses:
The system of claim 1, wherein end-user preferences are based on a developer specified schemta.
For claim 1 feature see claim 1 rejection, for rest of claim 12 feature see Yamanoue's FIG 2 and description in column 12, lines 25-28, "The user defined data 61 is data that the user can designate in relation to a data search, for each of the following items such as: an each user's search rule to generate queries: an alteration of the generated search rule; and frequency to conduct the search." -the user specified schema.

As Per Claim 13, Yamanoue discloses:
The system of claim 12, wherein information regarding end-user preferences and the developer schema are stored in one or more tables in the data stor~
For claim 12 feature see claim 12 rejection, for rest of claim 13 feature see Yamanoue's column 19, lines 9-11, "Then the query management means37 stores the filtering query matched with each user ID, which is stored in the user data DB 32, ":n a user ID table". - preference data stored in one or more ~bles in the data store.

Supplemental Box
In case **the space in any of the preceding boxes is not sufficient.**

As Per claim 18, Yamanoue discloses:
A method for employing preferences comprising: specifying user preferences based on a developer schema;
See claim 1 rejection.
storing the preferences in one or more tables in a data store;
See claim 13 rejection.
querying the tables in the data store upon occurrence of an event; producing a result table; and executing actions based on the data in the result table.
See Yamanoue's column 19, lines 9-19, "the query management means 37 stores the filtering query matched with each user ID, which is stored in the user data DB 32, in a user ID table (not illustrated) included in the query management means 37 ($54). The filtering query above is a query to acquire search results matched with a user ID from the search results made by the merged query which is merged in $53. That is to say, the filtering query can exclude search results matched with the dummy query from the search result of the merged query. So, the query management means 37 generates schedule data to conduct a data search by the queries registered in the query DB 36, from the search result of the merged query" - wherein each user's data or query is considered as an event. And further lines 30-35, "The query management means 37 registers queries which are not merged in $53 in the query DB 36, and also modifies/deletes queries registered in the query DB 36, when the query management means 37 determines that the queries are not used by a user possessing another user ID, by using a query management table (not illustrated) in the query management means 37 ($56)." -queries are executed aI~d the query results can be stored in user ID table or query management table (result table).

As Per claim 19, Yamanoue discloses:
The method of claim 18, wherein user preferences are specified by utilizing a one-at-a-time declarative programming model
For claim 18 feature see claim 18 rejection, for rest of claim 19 feature see Yamanoue's column 19, lines 21-23, "data about desirable time or duration of an update schedule (conduct of the search) designated by the user in the user defined data 61"; - the specified desirable time or duration of an update schedule designated by the user allows developers to specify one event against one preference, which is considered as using an 'one-at-a time' declarative model.

As Per claim 20, Yamanoue discloses:
The method of claim 19, wherein user preferences are specified using one or more On-event-If-Then statements and Boolean operators to specify conditions and actions.
For claim 19 feature see claim 19 rejection, for the 'on-event-If-Then' feature see see claim 2 (conditional statement) rejection.

As Per claim 21, Yamanoue discloses:
The method of claim 20, wherein querying the tables comprises executing query language statement.
For claim 20 feature see claim 20 rejection, for rest of claim 21 feature see Yamanoue's column 10, lines 48-50, "The search means 38" keeps track of the data server 2 capable of doing a data search by using queries generated in accordance with a specified search rule; provides queries specified by the query management means 37.for the data server 2; and gives instructions to the data server 2 to search data in accordance with the queries." - the queries specified by the query management means are inherently to be 'query language statements'.

As Per claim 23, Yamanoue discloses:
A computer readable medium having instructions stored thereon computer executable instructions for executing the method of claim 19.
For claim 19 feature see claim 19 rejection, for rest of claim 23 feature see Yamanoue's column 28, lines 34-36, "the operations described in the embodiments above can be realized by a computer program. The program is stored in a computer-readable storage medium. In the present invention, this storage medium may be a memory (not illustrated), for instance ROM itself, that is required when the search service center 3 operates, or may be a program storage medium that is readable by inserting into a program reader provided as an external storage device (not illustrated)."

Claims 14-17 lack of novelty under PCT Article 33(2) as being anticipated by US Patent No. 6,606,618 by Delo.
As Per claim 14, Delo discloses:
A method for application installation comprising: establishing a set of base tables; and updating the base tables with application data associated with an application being installed
Delo teaches a method for application installation, see Delo's column 2, lines 49-51, "The present invention fulfills the needs in the art by providing opfimizations for the process involved in the installation of a software product onto a target computer system", and column 6, lines 56-58, "the phrase 'software product' is meant to refer to an application program module (application) or a suite of application program modules." Further, in column 10, lines 7-10, "The present invention further extends the functionality of an installation database system by providing methods for efficiently managing and manipulating installation database tables." And updating the base tables with application data, see Delo's column 7, lines 15-28, the data elements stored in the database tables in FIG. 2A.

As Per claim 15, Delo discloses:
The method of claim 14, wherein the application is employs user defined preferences.
For claim 14 feature see claim 14 rejection, for rest of claim 15 feature see Delo's Fig. 2A and description in column 7, lines 15-28, "As shown in FIG. 2A, Feature_Table 205 stores a list of features, which are identified by Feature_ID 207. Each record in the Feature_Table 205 also includes a Feature_Name 209, an Attribute F A 211, an Attribute F B 212 and a Feature_Preference 213. The Component_Table

Supplemental Box
In case **the space in any of the preceding boxes is not sufficient.**

215 is constructed in the manner similar to the Feature_.Table 205, but stores records for components instead of features. Each record in the Component_Table 205 includes a Component_Name 219, an Attribute C A 221, an Attribute C B 223 and a Component Preference 223. The Feature_Comp_Table 225 associates Featitre_IDs 207 with Component_IDs 217. As can be seen, the data elements stored in the database tables of FIG. 2A comprise string data, each string having a variable length, and integer data." -application uses user defined preferences.

As Per claim 16, Delo discloses:
The method of claim 14, wherein application data includes application procedures that are stored as datn
For claim 14 feature see claim 14 rejection, Delo specifically mentions that "the
phrase 'software product' is meant to refer to an application program module
(application) or a suite of application program modules." - they are stored as data.

As Per claim 17, Delo discloses:
A computer readable medium having instructions stored thereon for carrying out the method of claim 14.
For claim 14 feature see claim 14 rejection, for rest of claim 17 see Delo's column 5, lines 38-40, The software product and installation database are generally provided via a computer-readable medium, such as a CD-ROM disk." Further in lines 64 to column 6, line 10, "the installation program module 137 may reside on the hard disk drive 127 of the target computer system and may operate in conjunction with the operating system ....In addition, the installation program
module may be provided along with the software product 136 and the installation database via a computer-readable medium, such as CD-ROM, or the like." - where the installation program, which includes instructions for carrying out the method of claim 14, is stored in a computer readable medium.

Claims 6, 8, 9, 10, and 12 lack an inventive step under PCT Article 33(3) as being obvious over US Patent No. 6,745,180 by Yamanoue in view of US 2003/0126136 by Omoigui.

As Per claim 6, Yamanoue discloses:
The system of claim 5, wherein the event source is a subscription servic~ For claim 5 feature see claim 5 rejection, Yamanoue .teaches all aspects of claim 6, but he does not disclose 'event source is a subscription service' explicitly, however, Omoigui teaches this feature in an analogous prior art; see Omoigui's paragraph [0254], "Network News Transfer Protocol (NNTP) ....
NNTP is designed so that news articles are stored in a central database allowing subscribers to select only those items they wish to read." And [0267], "this refers to all the data stored on users' local machines, in addition to user-specific data on an Agency server (e.g., subscribed server-side Agencies, server-side Favorite Agents, etc.)." and [0801], "allows users to browse, subscribe, and tmsubscribe to or from Agents on a given Agency that supports User State."
It would have been obvious to a person of ordinary skill in the art at the time of the invention was made to supplement Yamanoue's disclosure of the method of storing preference data in data store, and extract preference event from event source, by the event source is a subscription service taught by Omoigui. The modification would be obvious because one of ordinary skill in the artwould be motivated by delivering the user preference event data only to the interested parties, i.e. subscribed users/agents. (See Omoigui's paragraph [0267]).
As Per claim 8, Yamanoue discloses:
The system of claim 1, further comprising a conteaX analyzer to produce context data indicative of an end-users context at a given time and store the context data in the data stork
For claim 1 feature see claim 1 rejection, Yamanoue teaches all aspects of claim 8, but he does not disclose 'context analyzer at a given time' explicitly, however, Omoigui teaches this feature in an analogous prior art; see Omoigui's Abstract, "The system includes a first server component that is responsible for adding and maintaining domain-specific semantic information and a second server component that hosts semantic and other knowledge for use by the first server component that work together to provide context and time-sensitive semantic information retrieval services to clients operating a presentation platform via a communication medium."; see Omoigui's paragraph [0009], "Regardless of the search technique, the underlying organization of searchable information is index-driven rather than context-driven. The frequency or type of textual information associated the document determines the search results, as opposed to the attributes of the subject matter of the document and how those attributes relate to the user's context. (end-users context)"" Also see Omoigui's paragraph [0255], "The notification source (the client or server) stores information for the user and the Agent indicating the last time (stores data at a given time) the user acknowledged a notification for the Agent".
It would have been obvious to a person of ordinary skill in the art at the time of the invention was made to supplement Yamanoue's disclosure of the method of storing preference data in data store, and extract preference event from event source and store the data in data store, by the turning the available data into an end-users context, i.e. usable knowledge, taught by Omoigui. The modification would be obvious because one of ordinary skill in the art would be motivated by turning the data into meaningful context and efficient access for the users. (See Omoigui's paragraph [0006]).
As Per claim 9, Yamanoue discloses:
The system of claim 1, further comprising one or more APIs to interact with applications.
For claim 1 feature see claim 1 rejection, Yamanoue teaches all aspects of claim 9, but he does not disclose "one or more APIs to interact with applications' explicitly, however, Omoigui teaches this feature in an analogous prior art; see Omoigui's paragraph [0206], "Application Programming Interface (API). Defines how software programmers utilize a particular computer feature. APIs exist for

Supplemental Box
In case **the space in any of the preceding boxes is not sufficient.**

windowing systems, file systems, database systems, networking systems, and other systems."
It would have been obvious to a person of ordinary skill in the art at the time of the invention was made to supplement Yamanoue's disclosure of the method of storing preference data in data store, and extract preference event from event source and store the data in data store by using APIs to interact with applications taught by Omoigui. The modification would be obvious because one of ordinary skill in the art would be motivated by using APIs to capture input information, such as command parameters (See Omoigui's paragraph [0586]).
As Per claim 10, Yamanoue discloses:
The system of claim 1, wherein the compiler can compile and the execution engine can execute both heavyweight applications and lightweight preference applications.
For claim 1 feature see claim 1 rejection, Yamanoue teaches all aspects of claim 10, but he does not disclose 'execute both heavyweight applications and lightweight preference applications' explicitly, however, Omoigui teaches this feature in an analogous prior art; according to the description of the current application, paragraph [0011], "Heavyweight applications include those that are often run on high-end servers and require high-throughput and scalability, among other things. Lightweight applications are those that are often executed on smaller systems such as personal computers and require low-latency, a small database footprinL and small working set." Omoigui teaches the knowledge retrieval system runs on both the high-end servers and the 'smaller systems', see Omoigui's FIGURE 7, and paragraph [0008], "Information access further improved with the advent of the Internet, which connects a large number of computers across diverse geography to provide access to a vast body of information (heavyweight applications). The most wide spread method of providing information over the Internet is via the World Wide Web. The Web consists of a subset of the computers or Web servers connected to the Internet that typically run Hypertext Transfer Protocol (HTTP), File Transfer Protocol (FTP), GOPHER or other servers." And Omoigui's paragraph [0244], "Lightweight Directory Access Protocol (LDAP). Technology for accessing common directory information. LDAP has been embraced and implemented in most network-oriented middleware. As an open, vendor-neutral standard, LDAP provides an extendable architecture' for centralized storage and management of information (smaller systems such as personal computers and require low-latency, and small working set) that needs to be available for today's distributed systems and services. LDAP is currently supported in most network operating systems, groupware and even shrink-wrapped network applications."
It would have been obvious to a person of ordinary skill in the art at the time of the invention was made to supplement Yamanoue's disclosure of the method of extracting event from event source and storing preference data in data store by using both heavyweight and lightweight applications taught by Omoigui. The modification would be obvious because one of ordinary skill in the art would be motivated by presenting data to both low-end and high-end servers (See Omoigui's paragraph [0022]).
As Per claim 22, Yamanoue discloses:
The method of claim 19, wherein the developer schema is an XML schema. For claim 19 feature see claim 19 rejection, Yamanoue teaches all aspects of claim 22, but he does not disclose 'XML schema' explicitly, however, Omoigui teaches this feature in an analogous prior art, see Omoigui's paragraph [0008], "The most wide spread method of providing information over the Internet is via the World Wide Web. The Web consists of a subset of the computers or Web servers connected to the Internet that typically run Hypertext Transfer Protocol (HTTP), File Transfer Protocol 0~TP), GOPHER or other servers. Web servers host Web pages at Web sites. Web pages are encoded using one or more languages, such as the original Hypertext Markup Language (HTML) or the more current extensible Markup Language (XML) or the Standard Generic Markup Language (SGML)." It would have been obvious to a person of ordinary skill in the art at the time of the invention was made to supplement Yamanoue's disclosure of the method of storing preference data in data store, and extract preference event fxom event source, by the using XML taught by Omoigui. The modification would be obvious because one of ordinary skill in the art would be motivated by encapsulates structured, semantic queries in a flexible way (See Omoigui's paragraph [0202]).

NOTES TO FORM PCT/ISA/220 (continued)

The letter must indicate the differences between the claims as filed and the claims as amended. It must, in particular, indicate, in connection with each claim appearing in the international application (it being understood that identical indications concerning several claims may be grouped), whether

(i) the claim is unchanged;
(ii) the claim is cancelled;
(iii) the claim is new;
(iv) the claim replaces one or more claims as filed;
(v) the claim is the result of the division of a claim as filed.

The following examples illustrate the manner in which amendments must be explained in the accompanying letter:

1 [Where originally there were 48 claims and after amendment of some claims there are 51]:
"Claims 1 to 29, 31, 32, 34, 35, 37 to 48 replaced by amended claims bearing the same numbers, claims 30, 33 and 36 unchanged; new claims 49 to 51 added."

2 [Where originally there were 15 claims and after amendment of all claims there are 11]:
"Claims 1 to 15 replaced by amended claims 1 to 11."

3 [Where originally there were 14 claims and the amendments consist in cancelling some claims and in adding new claims]:
"Claims 1 to 6 and 14 unchanged; claims 7 to 13 cancelled; new claims 15, 16 and 17 added." or
"Claims 7 to 13 cancelled; new claims 15, 16 and 17 added; all other claims unchanged."

4 [Where various kinds of amendments are made]:
"Claims 1-10 unchanged; claims 11 to 13, 18 and 19 cancelled; claims 14, 15 and 16 replaced by amended claim 14; claim 17 subdivided into amended claims 15, 16 and 17; new claims 20 and 21 added."

"Statement under Article 19(1)" (Rule 46.4)

The amendments may be accompanied by a statement explaining the amendments and indicating any impact that such amendments might have on the description and the drawings (which cannot be amended under Article 19(1)).

The statement will be published with the international application and the amended claims.

It must be in the language in which the international application is to be published.

It must be brief, not exceeding 500 words if in English or if translated into English.

It should not be confused with and does not replace the letter indicating the differences between the claims as filed and as amended. It must be filed on a separate sheet and must be identified as such by a heading, preferably by using the words "Statement under Article 19(1)."

It may not contain any disparaging comments on the international search report or the relevance of citations contained in that report. Reference to citations, relevant to a given claim, contained in the international search report may be made only in connection with an amendment of that claim.

Consequence if a demand for international preliminary examination has already been filed

If, at the time of filing any amendments and any accompanying statement, under Article 19, a demand for international preliminary examination has already been submitted, the applicant must preferably, at the time of filing the amendments (and any statement) with the International Bureau, also file with the International Preliminary Examining Authority a copy of such amendments (and of any statement) and, where required, a translation of such amendments for the procedure before that Authority (see Rules 55.3(a) and 62.2, first sentence). For further information, see the Notes to the demand form (PCT/IPEA/401).

If a demand for international preliminary examination is made, the written opinion of the International Searching Authority will, except in certain cases where the International Preliminary Examining Authority did not act as International Searching Authority and where it has notified the International Bureau under Rule 66.1*bis*(b), be considered to be a written opinion of the International Preliminary Examining Authority. If a demand is made, the applicant may submit to the International Preliminary Examining Authority a reply to the written opinion together, where appropriate, with amendments before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later (Rule 43*bis*.1(c)).

Consequence with regard to translation of the international application for entry into the national phase

The applicant's attention is drawn to the fact that, upon entry into the national phase, a translation of the claims as amended under Article 19 may have to be furnished to the designated/elected Offices, instead of, or in addition to, the translation of the claims as filed.

For further details on the requirements of each designated/elected Office, see the *PCT Applicant's Guide*, Volume II.

NOTES TO FORM PCT/ISA/220

These Notes are intended to give the basic instructions concerning the filing of amendments under Article 19. The Notes are based on the requirements of the Patent Cooperation Treaty, the Regulations and the Administrative Instructions under that Treaty. In case of discrepancy between these Notes and those requirements, the latter are applicable. For more detailed information, see also the *PCT Applicant's Guide*, a publication of WIPO.

In these Notes, "Article," "Rule" and "Section" refer to the provisions of the PCT, the PCT Regulations and the PCT Administrative Instructions, respectively

INSTRUCTIONS CONCERNING AMENDMENTS UNDER ARTICLE 19

The applicant has, after having received the international search report and the written opinion of the International Searching Authority, one opportunity to amend the claims of the international application. It should however be emphasized that, since all parts of the international application (claims, description and drawings) may be amended during the international preliminary examination procedure, there is usually no need to file amendments of the claims under Article 19 except where, e.g. the applicant wants the latter to be published for the purposes of provisional protection or has another reason for amending the claims before international publication. Furthermore, it should be emphasized that provisional protection is available in some States only (see *PCT Applicant's Guide*, Volume I/A, Annexes B1 and B2).

The attention of the applicant is drawn to the fact that amendments to the claims under Article 19 are not allowed where the International Searching Authority has declared, under Article 17(2), that no international search report would be established (see *PCT Applicant's Guide*, Volume I/A, paragraph 296).

What parts of the international application may be amended ?

Under Article 19, only the claims may be amended.

During the international phase, the claims may also be amended (or further amended) under Article 34 before the International Preliminary Examining Authority. The description and drawings may only be amended under Article 34 before the International Preliminary Examining Authority.

Upon entry into the national phase, all parts of the international application may be amended under Article 28 or, where applicable, Article 41.

When ? Within 2 months from the date of transmittal of the international search report or 16 months from the priority date, whichever time limit expires later. It should be noted, however, that the amendments will be considered as having been received on time if they are received by the International Bureau after the expiration of the applicable time limit but before the completion of the technical preparations for international publication (Rule 46.1).

Where not to file the amendments ?

The amendments may only be filed with the International Bureau and not with the receiving Office or the International Searching Authority (Rule 46.2).

Where a demand for international preliminary examination has been/is filed, see below.

How ? Either by cancelling one or more entire claims, by adding one or more new claims or by amending the text of one or more of the claims as filed.

A replacement sheet must be submitted for each sheet of the claims which, on account of an amendment or amendments, differs from the sheet originally filed.

All the claims appearing on a replacement sheet must be numbered in Arabic numerals. Where a claim is cancelled, no renumbering of the other claims is required. In all cases where claims are renumbered, they must be renumbered consecutively (Section 205(b)).

The amendments must be made in the language in which the international application is to be published.

What documents must/may accompany the amendments ?

Letter (Section 205(b)):

The amendments must be submitted with a letter.

The letter will not be published with the international application and the amended claims. It should not be confused with the "Statement under Article 19(1)" (see below, under "Statement under Article 19(1)").

The letter must be in English or French, at the choice of the applicant. However, if the language of the international application is English, the letter must be in English; if the language of the international application is French, the letter must be in French.

US006745180B2

(12) **United States Patent**
Yamanoue

(10) **Patent No.: US 6,745,180 B2**
(45) **Date of Patent: Jun. 1, 2004**

(54) **DATA SUPPLY CONTROLLING DEVICE, METHOD, AND STORAGE MEDIUM WHICH FACILITIES INFORMATION SEARCHING BY USER**

(75) Inventor: **Masafumi Yamanoue**, Yamatokoriyama (JP)

(73) Assignee: **Sharp Kabushiki Kaisha**, Osaka (JP)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 316 days.

(21) Appl. No.: **09/976,048**

(22) Filed: **Oct. 15, 2001**

(65) **Prior Publication Data**

US 2002/0103806 A1 Aug. 1, 2002

(30) **Foreign Application Priority Data**

Oct. 18, 2000 (JP) 2000-318318

(51) **Int. Cl.**[7] **G06F 17/30**

(52) **U.S. Cl.** **707/3**; 707/2; 707/4; 707/5; 707/6

(58) **Field of Search** 707/2, 3, 4, 5, 707/6

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,519,859	A *	5/1996	Grace	707/3
5,632,015	A *	5/1997	Zimowski et al.	707/3
5,864,841	A *	1/1999	Agrawal et al.	707/2
6,148,296	A *	11/2000	Tabbara	707/3
6,366,910	B1 *	4/2002	Rajaraman et al.	707/5
2001/0047355	A1 *	11/2001	Anwar	707/5
2002/0194167	A1 *	12/2002	Bakalash et al.	707/3
2003/0163466	A1 *	8/2003	Rajaraman et al.	707/6

FOREIGN PATENT DOCUMENTS

EP	0314279 A2 *	5/1989	G06F/15/40
JP	2976219 B2	9/1999	
JP	11-259512 A	9/1999	

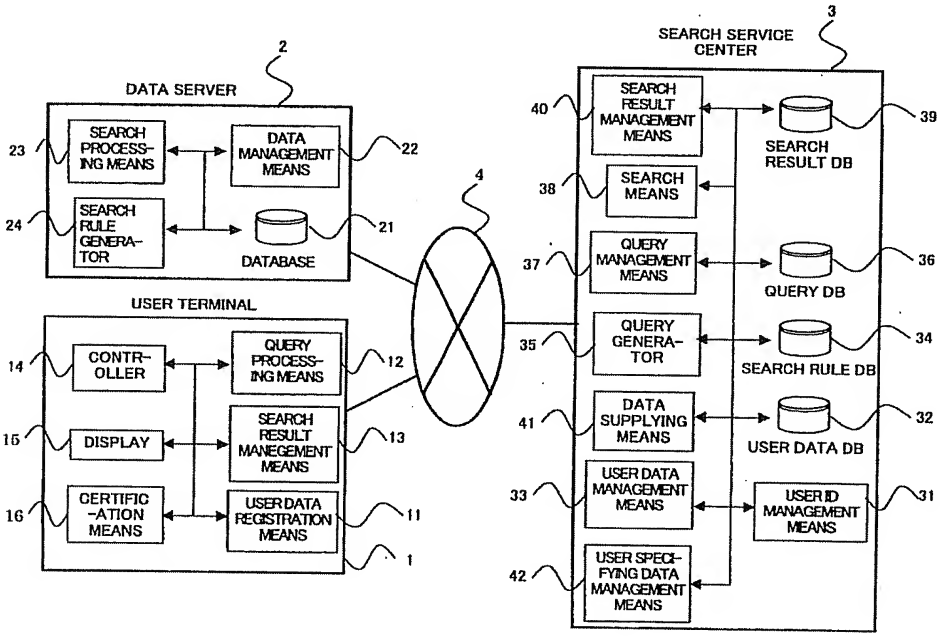
* cited by examiner

Primary Examiner—Frantz Coby
(74) *Attorney, Agent, or Firm*—Nixon & Vanderhye, P.C.

(57) **ABSTRACT**

A data supply controlling device comprises a data base for user data which stores user data matched with each user. The data base of user data can be queried in accordance with the user data so that a data server performs a search according to the query and stores the search results in a search result data base. User ID management can be performed for separately managing the identifying data to identify each user and user specifying data to specify each user. A search result matched with the identifying data of the user from the search results stored in the search result data base is provided for a user terminal. Once the user data is stored, it is not required to carry out further communications to input the user data. Privacy protection of the user and improvement of communications efficiency are both fulfilled in the system to search for information in accordance with user data.

39 Claims, 16 Drawing Sheets



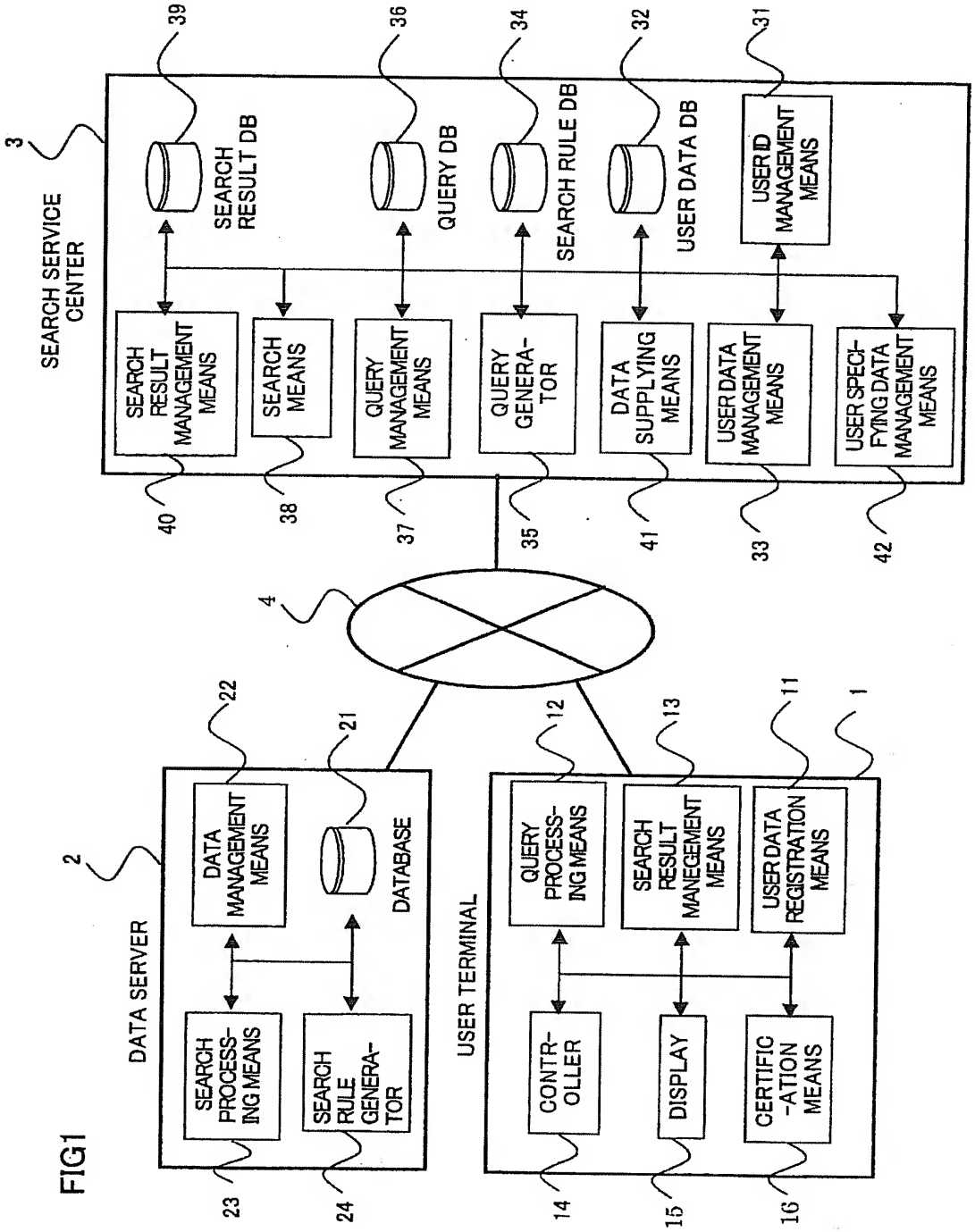

SEARCH SERVICE CENTER
3
SEARCH RESULT DB
39
QUERY DB
36
SEARCH RULE DB
34
USER DATA DB
32
USER ID MANAGEMENT MEANS
31
SEARCH RESULT MANAGEMENT MEANS
40
SEARCH MEANS
38
QUERY MANAGEMENT MEANS
37
QUERY GENERA-TOR
35
DATA SUPPLYING MEANS
41
USER DATA MANAGEMENT MEANS
33
USER SPECI-FYING DATA MANAGEMENT MEANS
42
4
DATA SERVER
2
DATA MANAGEMENT MEANS
22
DATABASE
21
SEARCH PROCESS-ING MEANS
23
SEARCH RULE GENERA-TOR
24
USER TERMINAL
1
QUERY PROCESS-ING MEANS
12
SEARCH RESULT MANEGEMENT MEANS
13
USER DATA REGISTRATION MEANS
11
CONTR-OLLER
14
DISPLAY
15
CERTIFIC-ATION MEANS
16

FIG1

FIG2

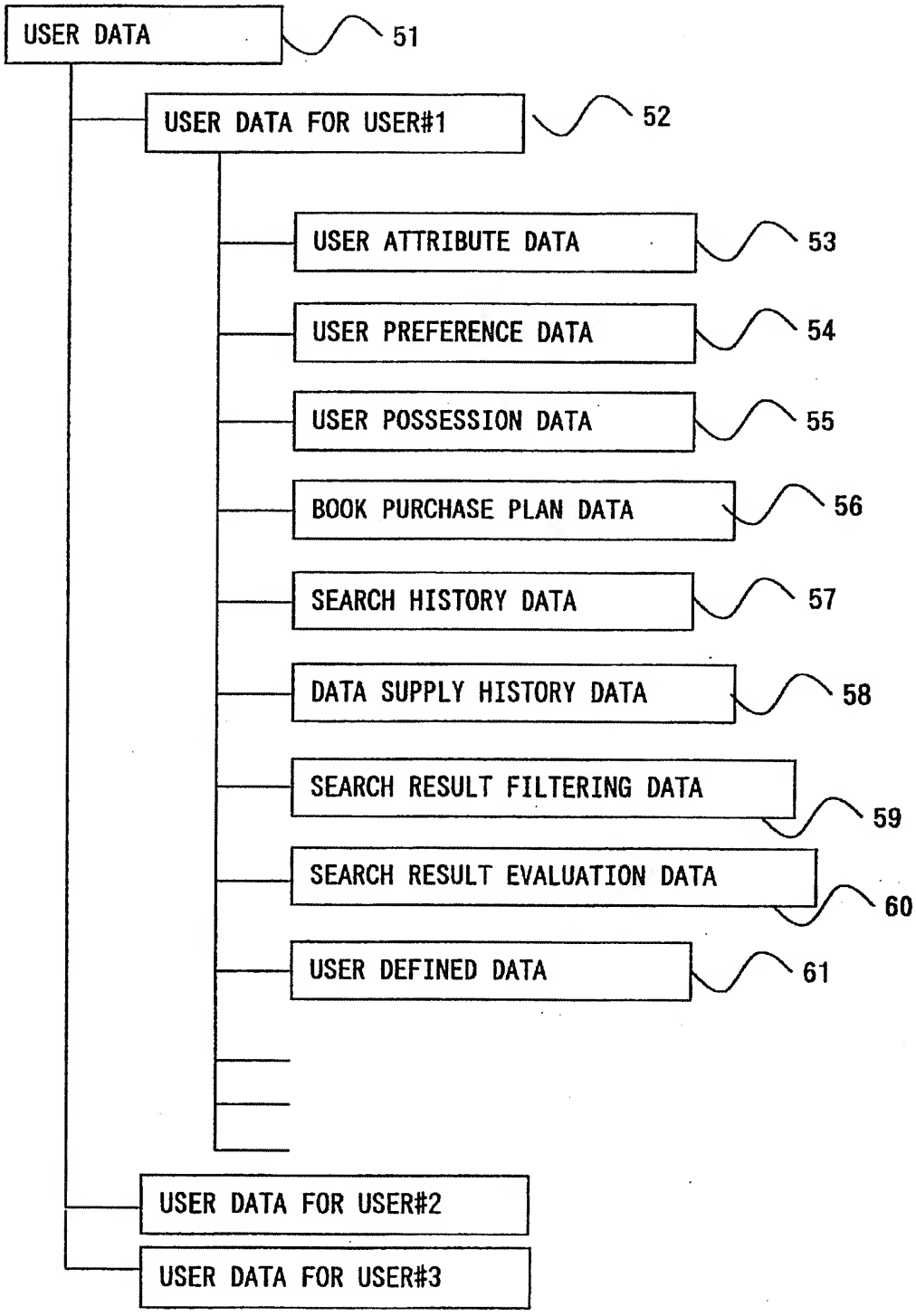
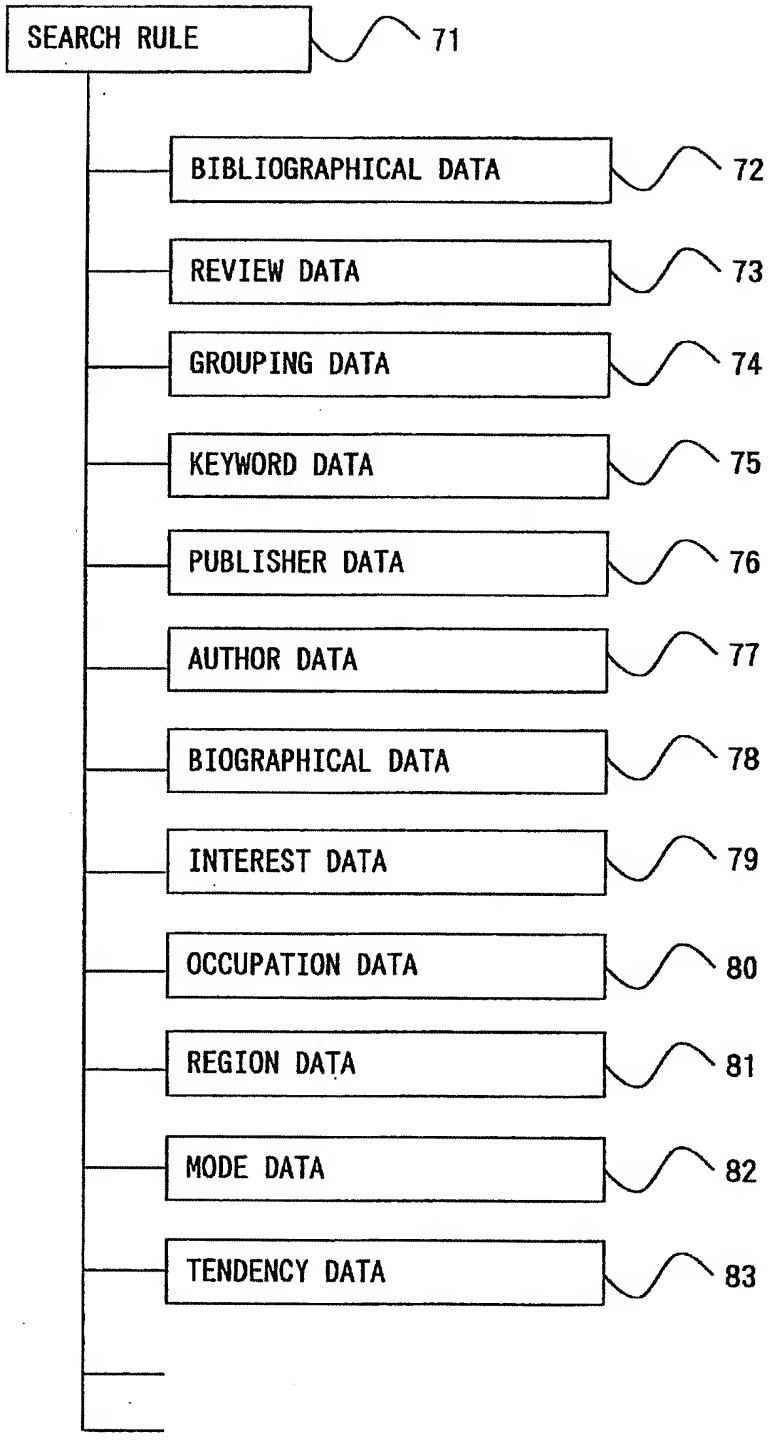

USER DATA
51
USER DATA FOR USER#1
52
USER ATTRIBUTE DATA
53
USER PREFERENCE DATA
54
USER POSSESSION DATA
55
BOOK PURCHASE PLAN DATA
56
SEARCH HISTORY DATA
57
DATA SUPPLY HISTORY DATA
58
SEARCH RESULT FILTERING DATA
59
SEARCH RESULT EVALUATION DATA
60
USER DEFINED DATA
61
USER DATA FOR USER#2
USER DATA FOR USER#3

FIG3


SEARCH RULE
71
BIBLIOGRAPHICAL DATA
72
REVIEW DATA
73
GROUPING DATA
74
KEYWORD DATA
75
PUBLISHER DATA
76
AUTHOR DATA
77
BIOGRAPHICAL DATA
78
INTEREST DATA
79
OCCUPATION DATA
80
REGION DATA
81
MODE DATA
82
TENDENCY DATA
83

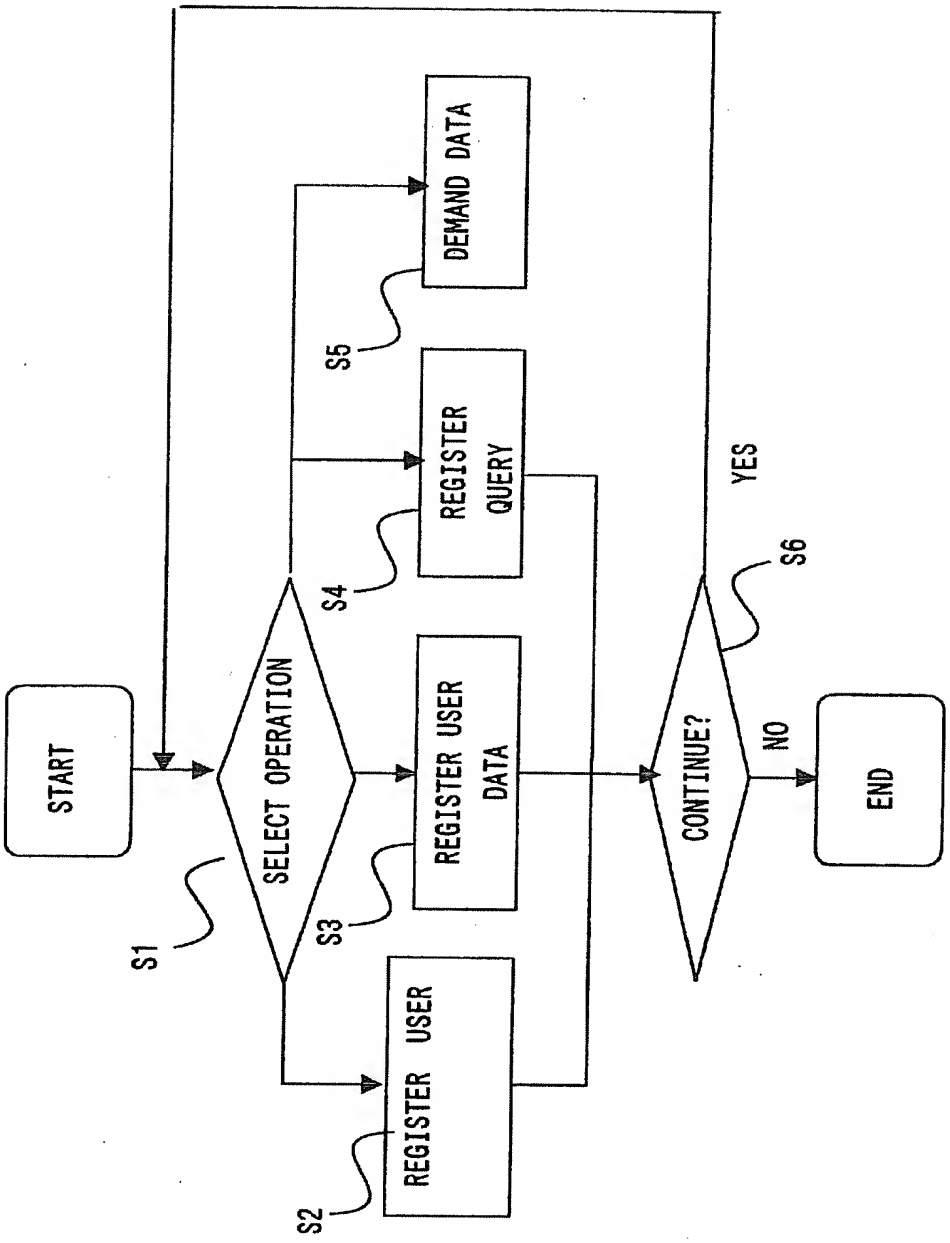

START
SELECT OPERATION
S1
REGISTER USER
S2
REGISTER USER DATA
S3
REGISTER QUERY
S4
DEMAND DATA
S5
CONTINUE?
S6
YES
NO
END

FIG4

FIG5

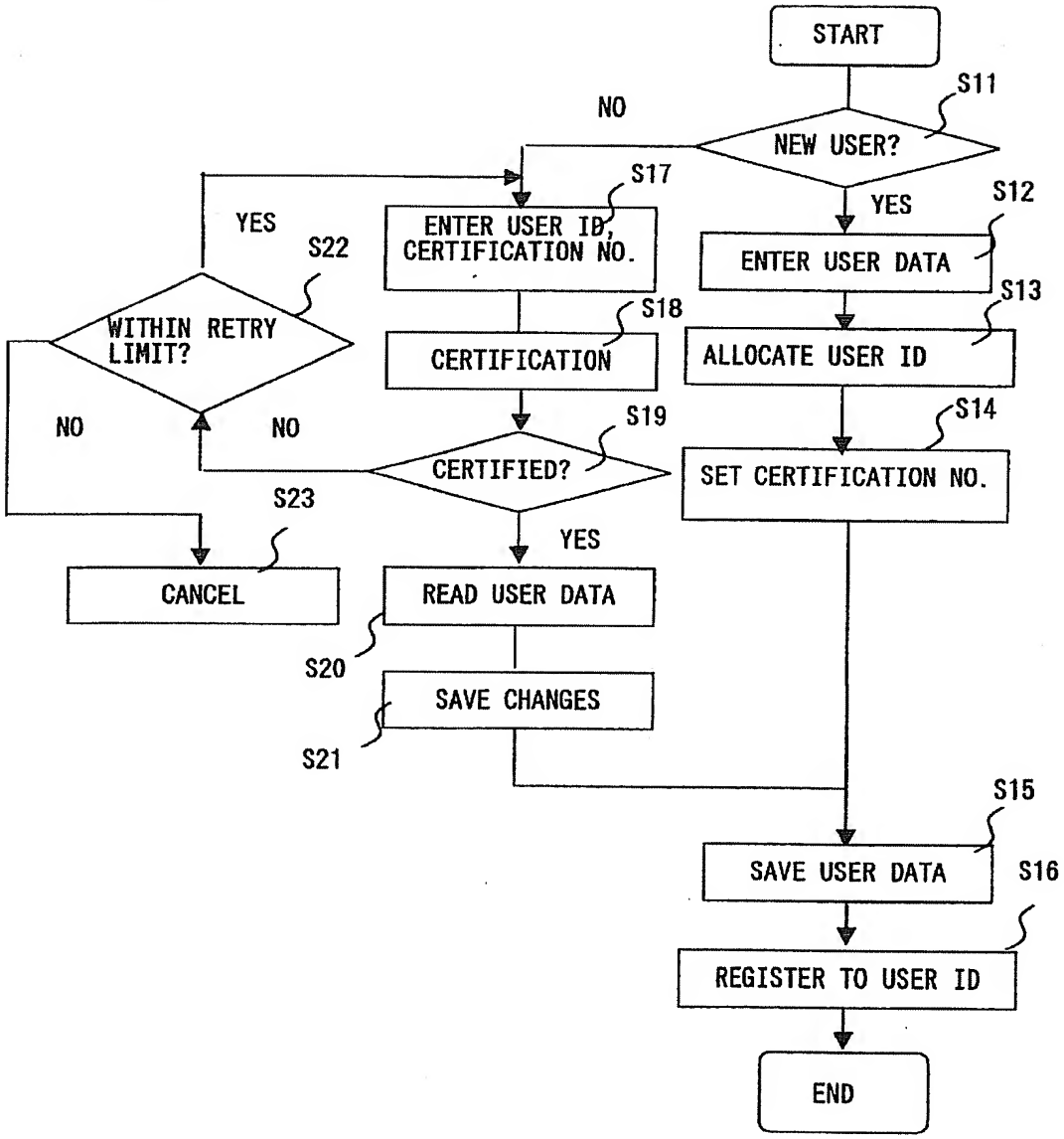

START
S11
NEW USER?
NO
YES
S12
ENTER USER DATA
S13
ALLOCATE USER ID
S14
SET CERTIFICATION NO.
S15
SAVE USER DATA
S16
REGISTER TO USER ID
END
S17
ENTER USER ID, CERTIFICATION NO.
S18
CERTIFICATION
S19
CERTIFIED?
NO
YES
S22
WITHIN RETRY LIMIT?
NO
YES
S23
CANCEL
S20
READ USER DATA
S21
SAVE CHANGES

FIG6

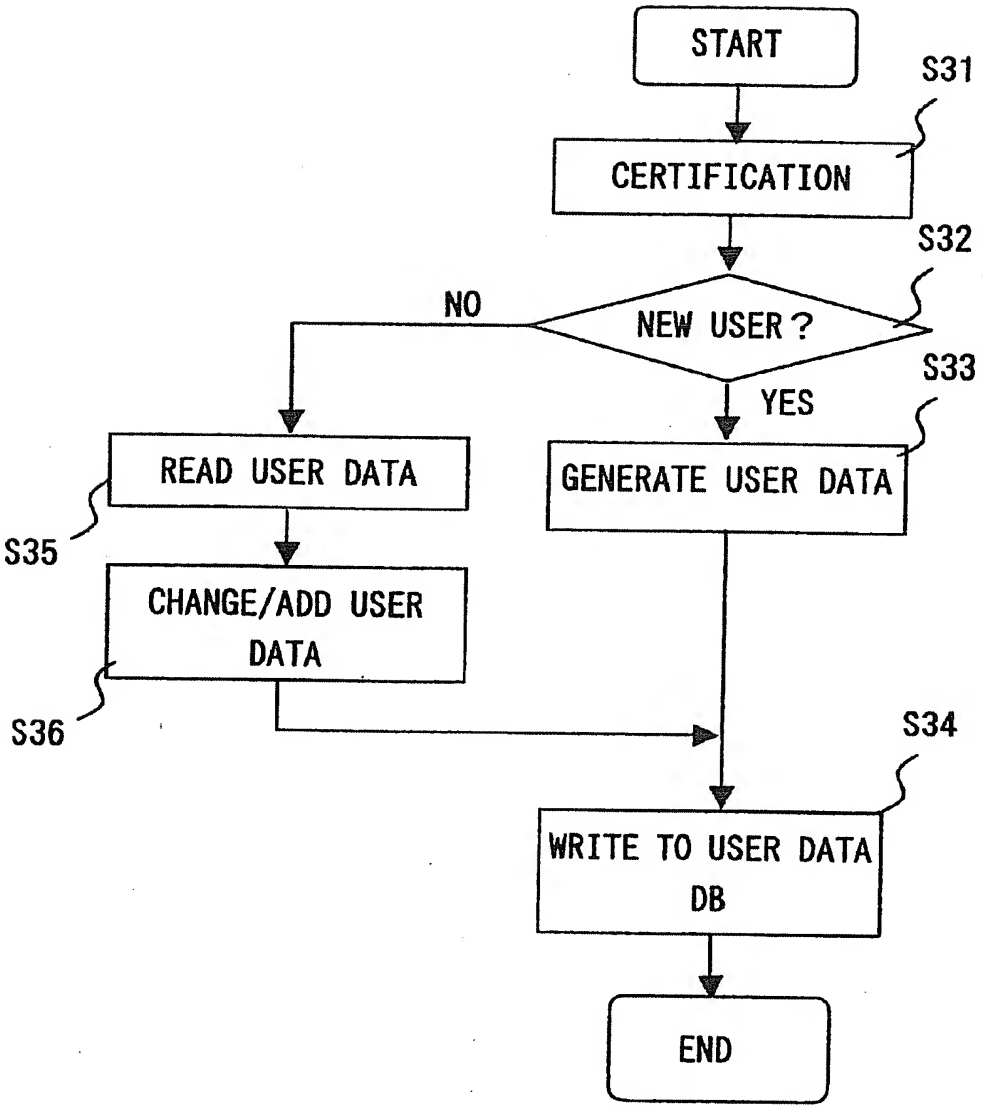

START
S31
CERTIFICATION
S32
NEW USER?
NO
YES
S33
READ USER DATA
GENERATE USER DATA
S35
CHANGE/ADD USER DATA
S36
S34
WRITE TO USER DATA DB
END

FIG7

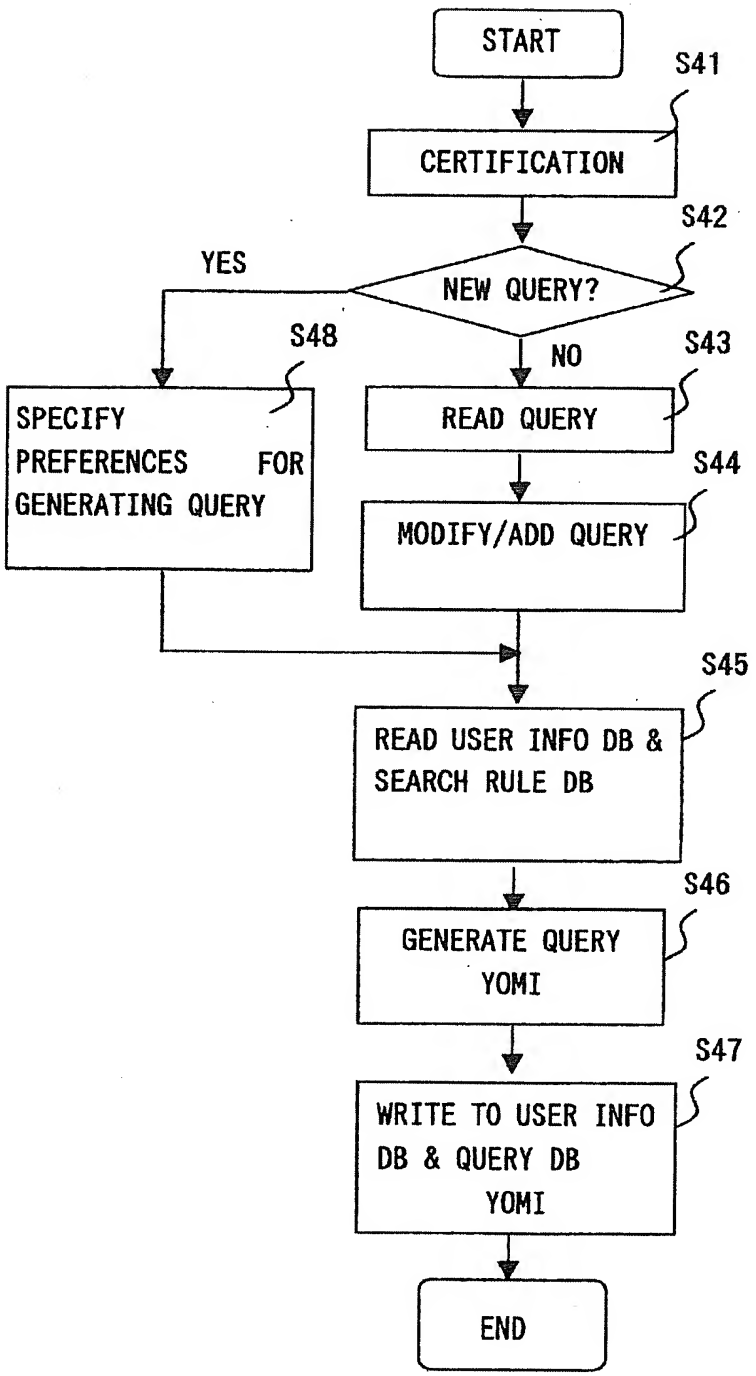

START
S41
CERTIFICATION
S42
NEW QUERY?
YES
NO
S48
SPECIFY PREFERENCES FOR GENERATING QUERY
S43
READ QUERY
S44
MODIFY/ADD QUERY
S45
READ USER INFO DB & SEARCH RULE DB
S46
GENERATE QUERY YOMI
S47
WRITE TO USER INFO DB & QUERY DB YOMI
END

FIG8

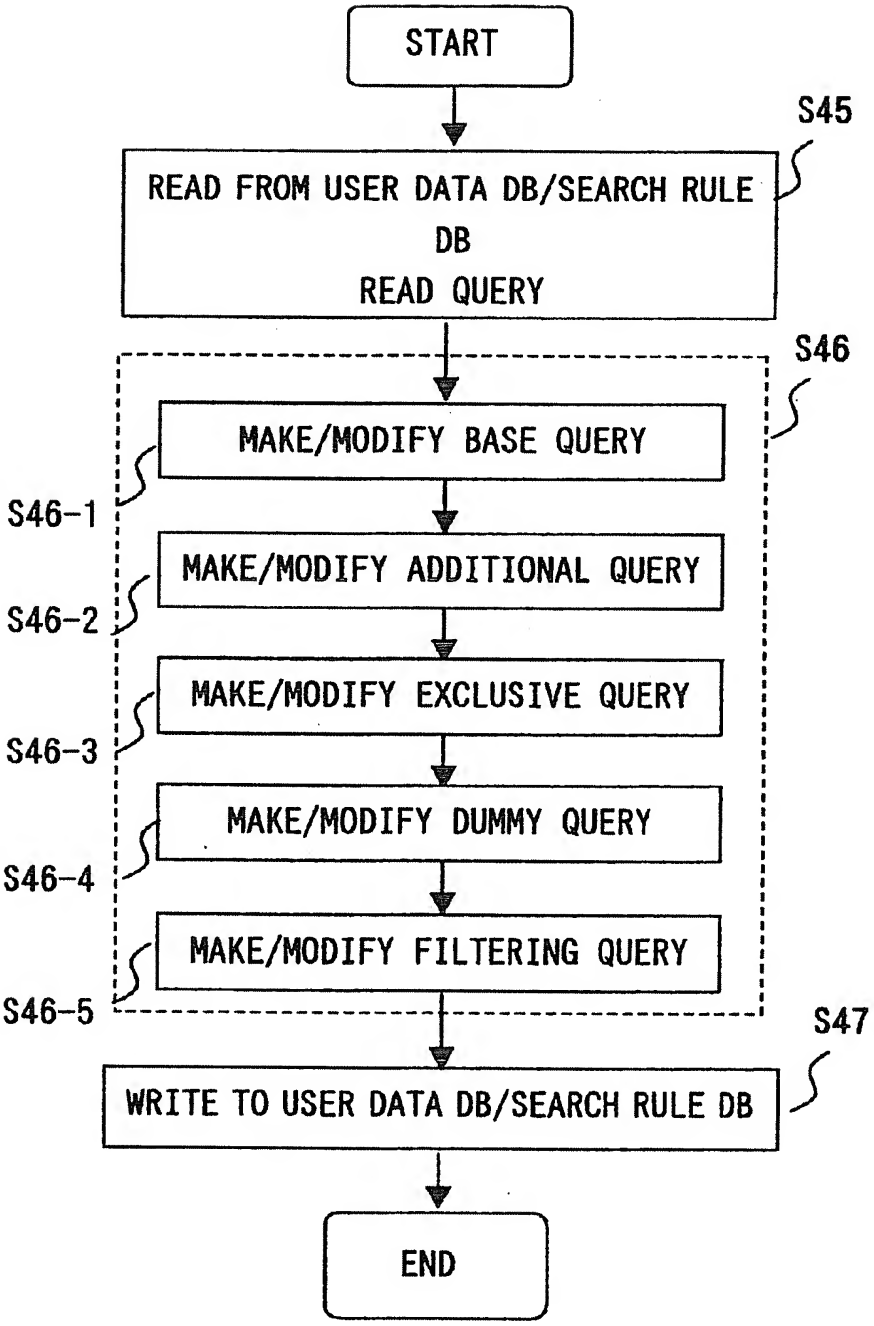
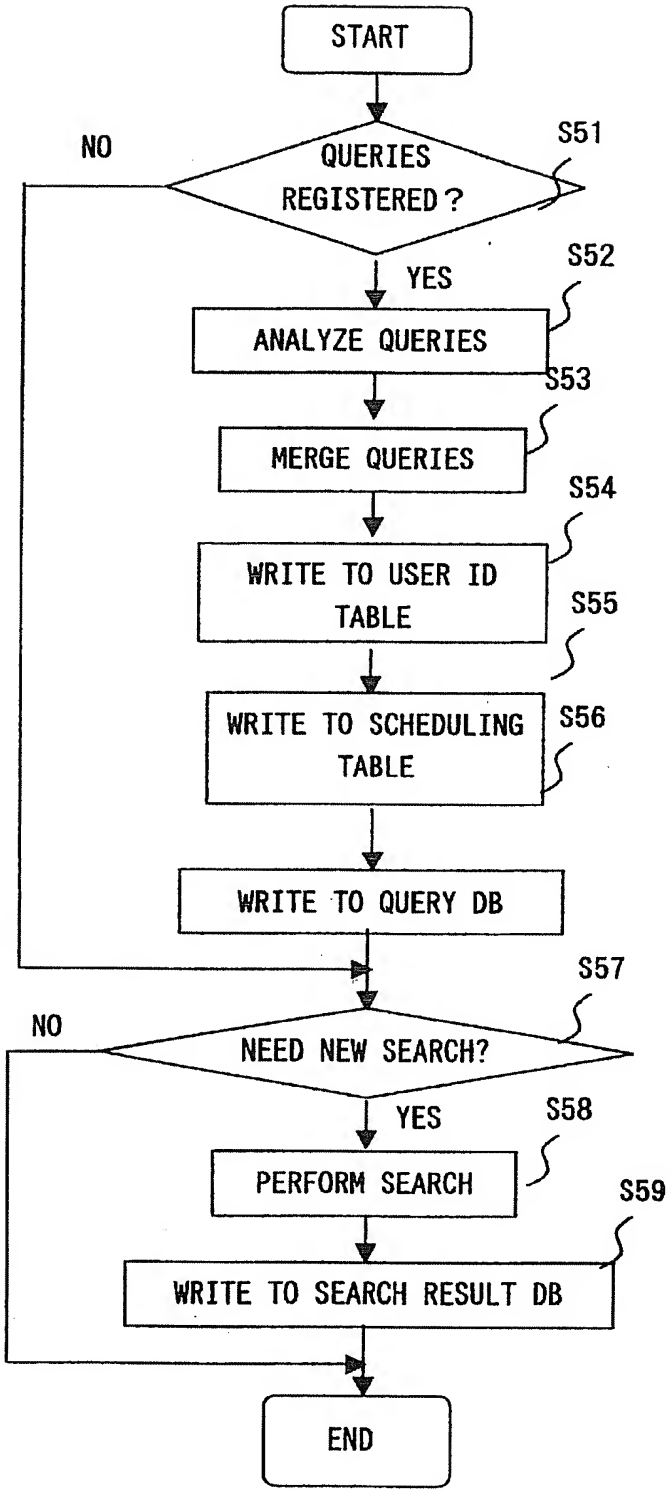

START
S45
READ FROM USER DATA DB/SEARCH RULE DB
READ QUERY
S46
MAKE/MODIFY BASE QUERY
S46-1
MAKE/MODIFY ADDITIONAL QUERY
S46-2
MAKE/MODIFY EXCLUSIVE QUERY
S46-3
MAKE/MODIFY DUMMY QUERY
S46-4
MAKE/MODIFY FILTERING QUERY
S46-5
S47
WRITE TO USER DATA DB/SEARCH RULE DB
END

FIG9


START
QUERIES REGISTERED ?
S51
NO
YES
S52
ANALYZE QUERIES
S53
MERGE QUERIES
S54
WRITE TO USER ID TABLE
S55
WRITE TO SCHEDULING TABLE
S56
WRITE TO QUERY DB
S57
NO
NEED NEW SEARCH?
YES
S58
PERFORM SEARCH
S59
WRITE TO SEARCH RESULT DB
END

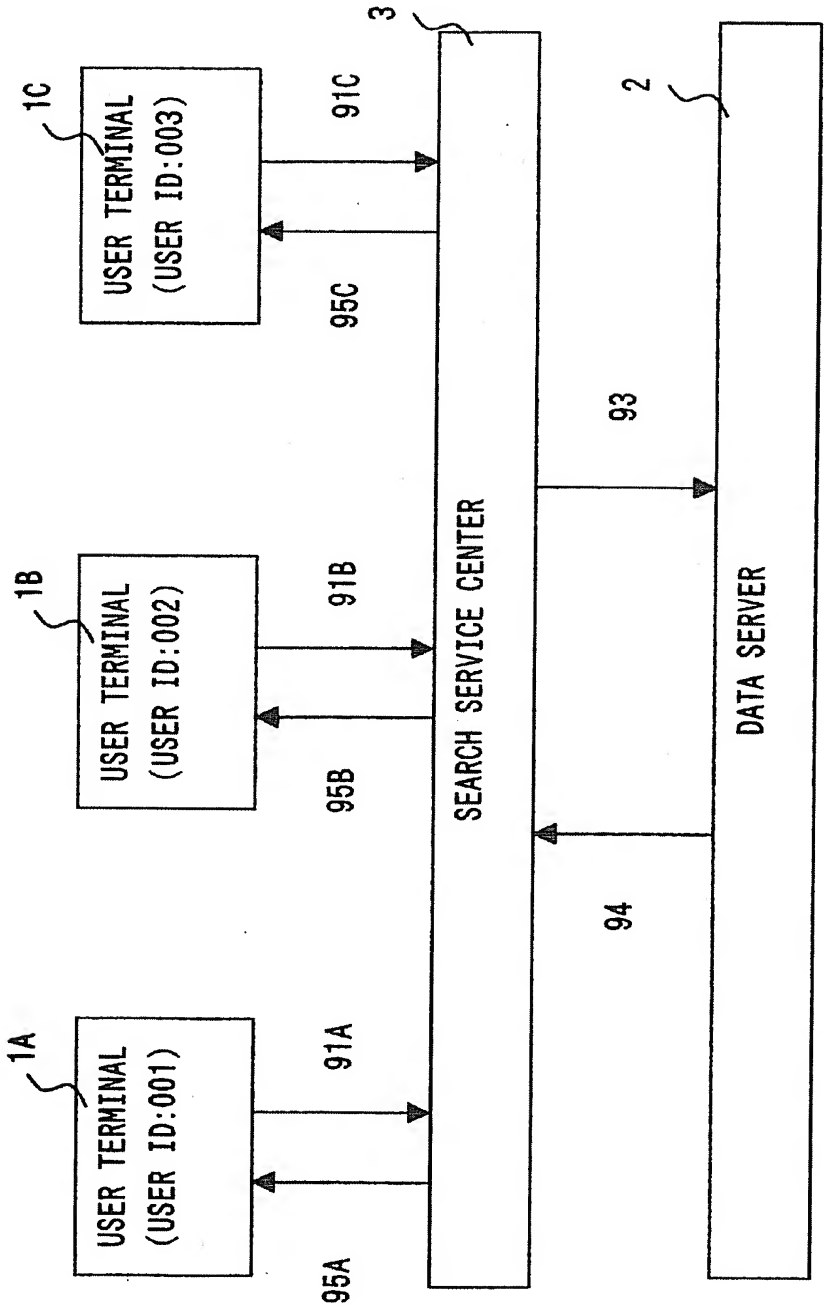

1A
USER TERMINAL
(USER ID:001)
1B
USER TERMINAL
(USER ID:002)
1C
USER TERMINAL
(USER ID:003)
91A
91B
91C
95A
95B
95C
3
SEARCH SERVICE CENTER
93
94
2
DATA SERVER

FIG10

FIG11

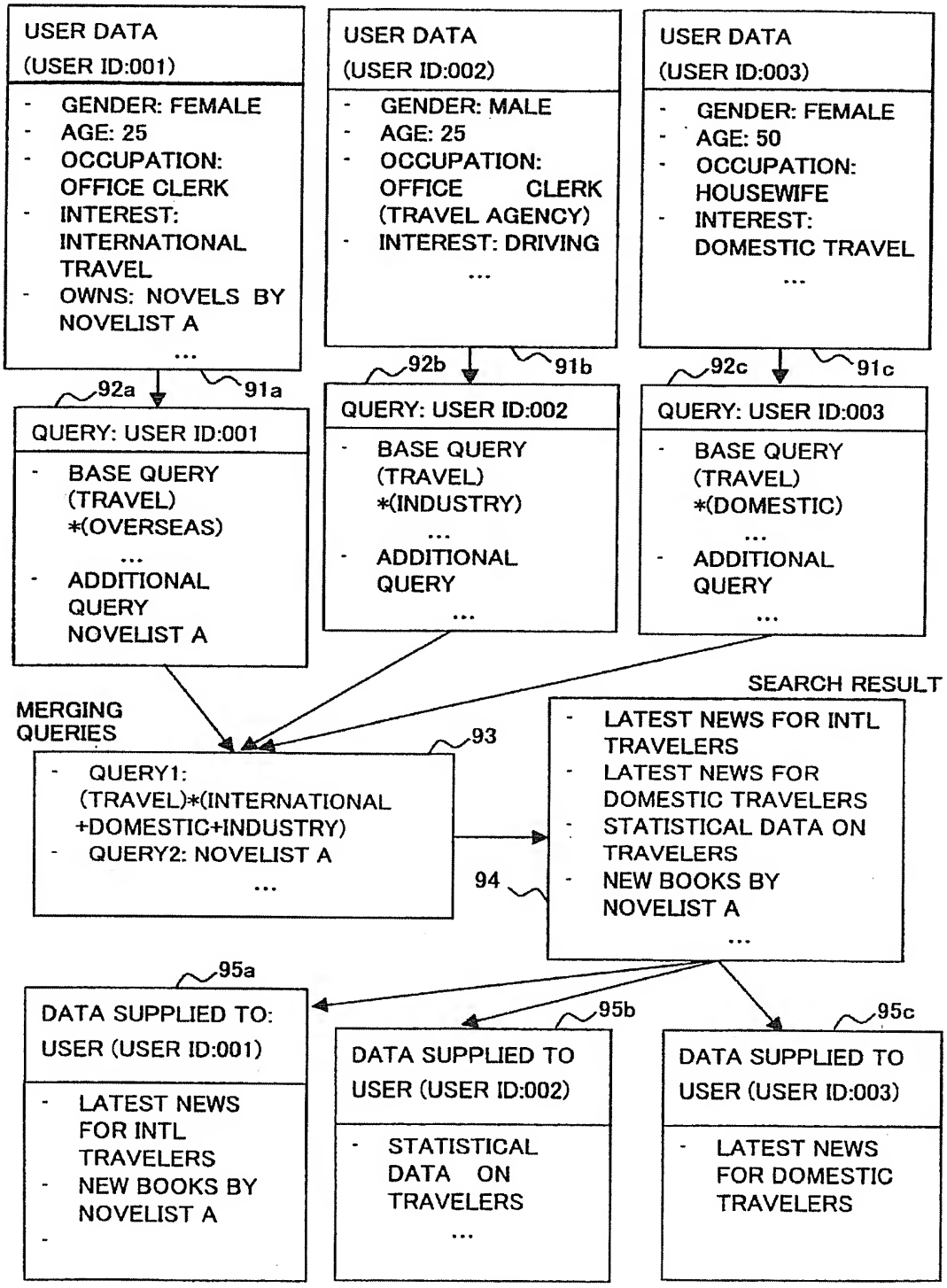
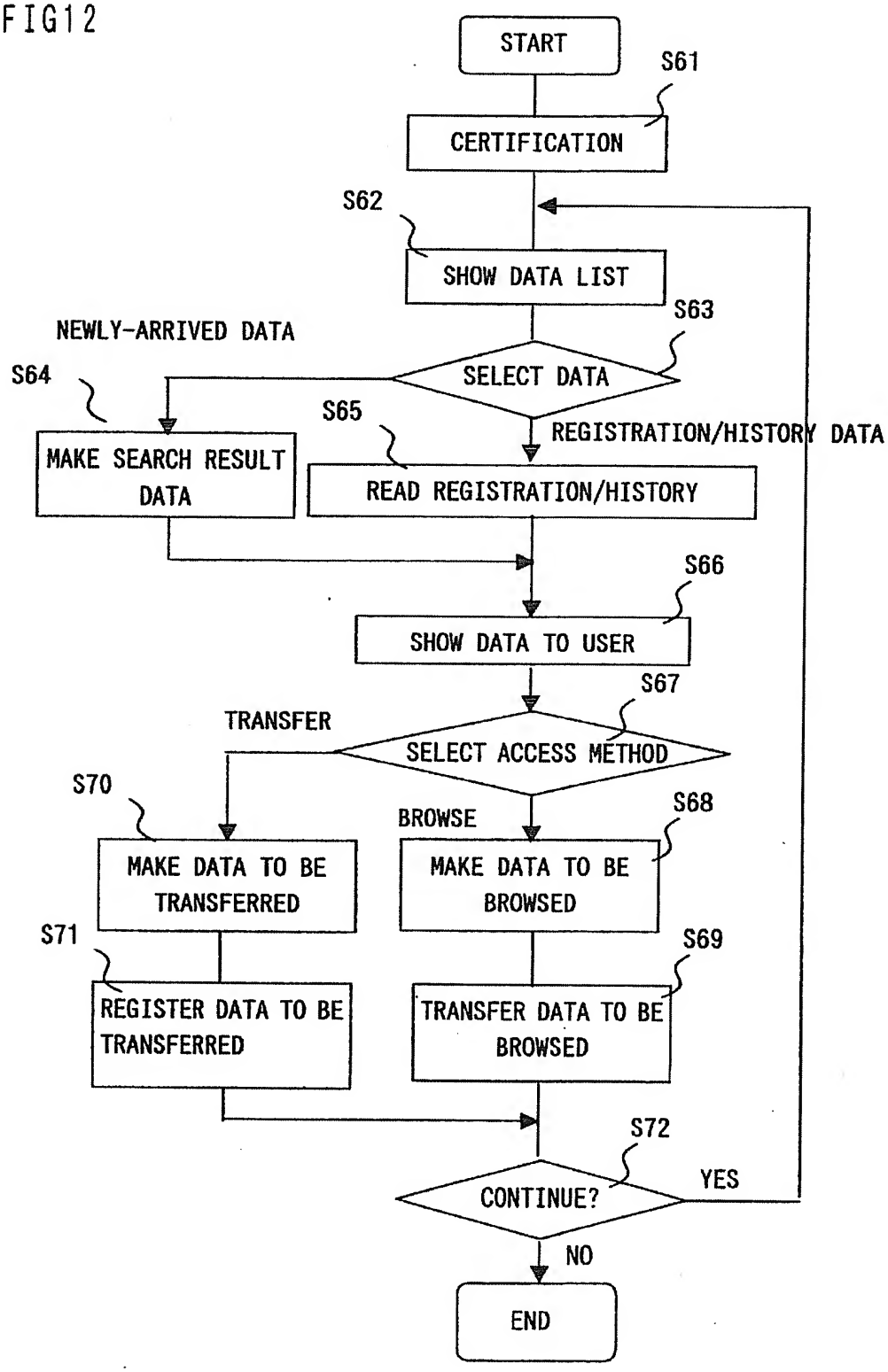

USER DATA
(USER ID:001)
- GENDER: FEMALE
- AGE: 25
- OCCUPATION: OFFICE CLERK
- INTEREST: INTERNATIONAL TRAVEL
- OWNS: NOVELS BY NOVELIST A
...
USER DATA
(USER ID:002)
- GENDER: MALE
- AGE: 25
- OCCUPATION: OFFICE CLERK (TRAVEL AGENCY)
- INTEREST: DRIVING
...
USER DATA
(USER ID:003)
- GENDER: FEMALE
- AGE: 50
- OCCUPATION: HOUSEWIFE
- INTEREST: DOMESTIC TRAVEL
...
91a
92a
91b
92b
91c
92c
QUERY: USER ID:001
- BASE QUERY (TRAVEL) *(OVERSEAS)
...
- ADDITIONAL QUERY NOVELIST A
QUERY: USER ID:002
- BASE QUERY (TRAVEL) *(INDUSTRY)
...
- ADDITIONAL QUERY
...
QUERY: USER ID:003
- BASE QUERY (TRAVEL) *(DOMESTIC)
...
- ADDITIONAL QUERY
...
MERGING QUERIES
93
- QUERY1: (TRAVEL)*(INTERNATIONAL +DOMESTIC+INDUSTRY)
- QUERY2: NOVELIST A
...
SEARCH RESULT
- LATEST NEWS FOR INTL TRAVELERS
- LATEST NEWS FOR DOMESTIC TRAVELERS
- STATISTICAL DATA ON TRAVELERS
- NEW BOOKS BY NOVELIST A
...
94
95a
95b
95c
DATA SUPPLIED TO: USER (USER ID:001)
- LATEST NEWS FOR INTL TRAVELERS
- NEW BOOKS BY NOVELIST A
-
DATA SUPPLIED TO USER (USER ID:002)
- STATISTICAL DATA ON TRAVELERS
...
DATA SUPPLIED TO USER (USER ID:003)
- LATEST NEWS FOR DOMESTIC TRAVELERS

FIG12


START
S61
CERTIFICATION
S62
SHOW DATA LIST
S63
NEWLY-ARRIVED DATA
SELECT DATA
S64
S65
REGISTRATION/HISTORY DATA
MAKE SEARCH RESULT DATA
READ REGISTRATION/HISTORY
S66
SHOW DATA TO USER
S67
TRANSFER
SELECT ACCESS METHOD
S70
BROWSE
S68
MAKE DATA TO BE TRANSFERRED
MAKE DATA TO BE BROWSED
S71
S69
REGISTER DATA TO BE TRANSFERRED
TRANSFER DATA TO BE BROWSED
S72
CONTINUE?
YES
NO
END

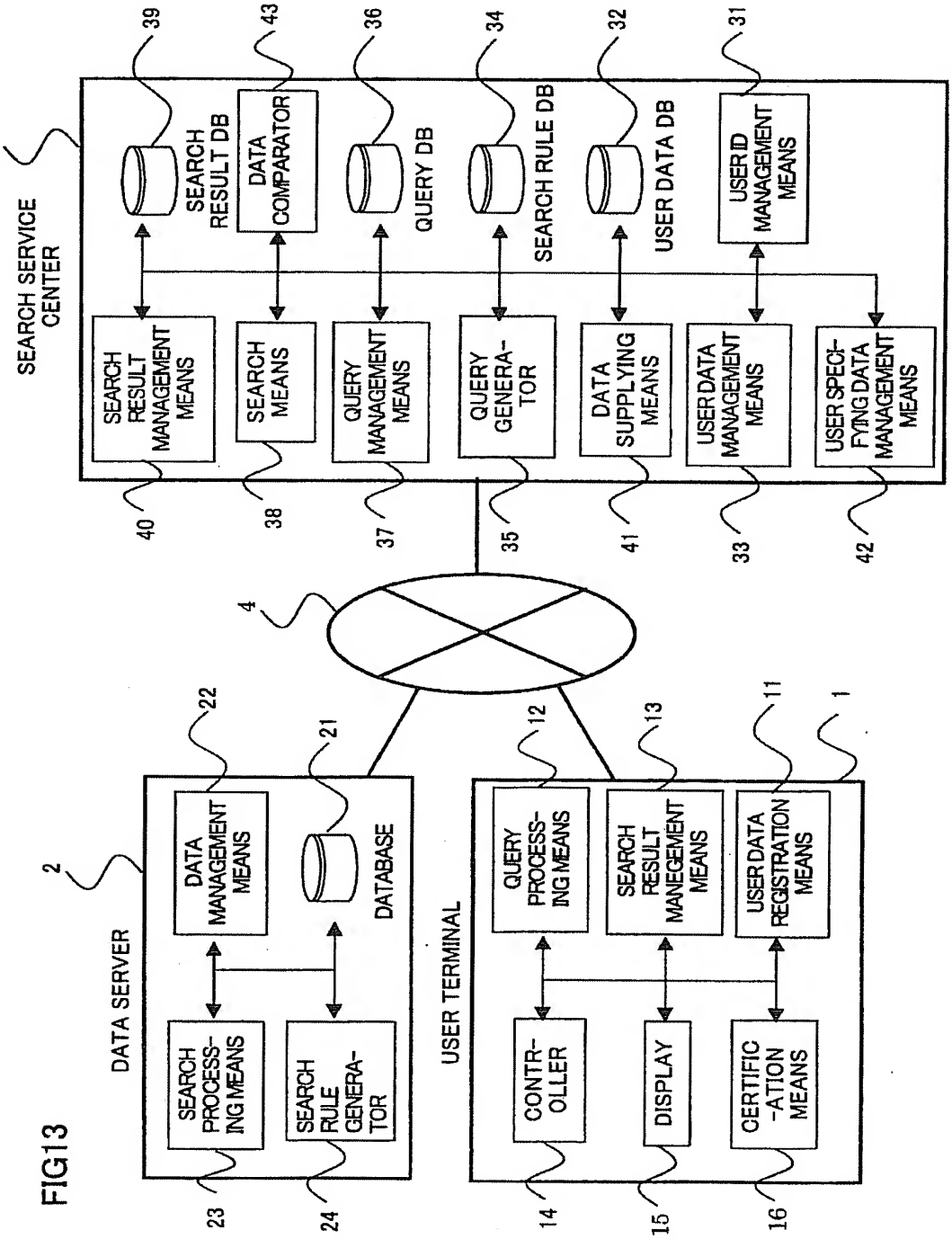

SEARCH SERVICE CENTER
3
SEARCH RESULT DB
DATA COMPARATOR
QUERY DB
SEARCH RULE DB
USER DATA DB
USER ID MANAGEMENT MEANS
SEARCH RESULT MANAGEMENT MEANS
SEARCH MEANS
QUERY MANAGEMENT MEANS
QUERY GENERA-TOR
DATA SUPPLYING MEANS
USER DATA MANAGEMENT MEANS
USER SPECI-FYING DATA MANAGEMENT MEANS
39
43
36
34
32
31
40
38
37
35
41
33
42
4
DATA SERVER
2
DATA MANAGEMENT MEANS
DATABASE
SEARCH PROCESS-ING MEANS
SEARCH RULE GENERA-TOR
22
21
23
24
USER TERMINAL
1
QUERY PROCESS-ING MEANS
SEARCH RESULT MANEGEMENT MEANS
USER DATA REGISTRATION MEANS
CONTR-OLLER
DISPLAY
CERTIFIC-ATION MEANS
12
13
11
14
15
16

FIG13

FIG14

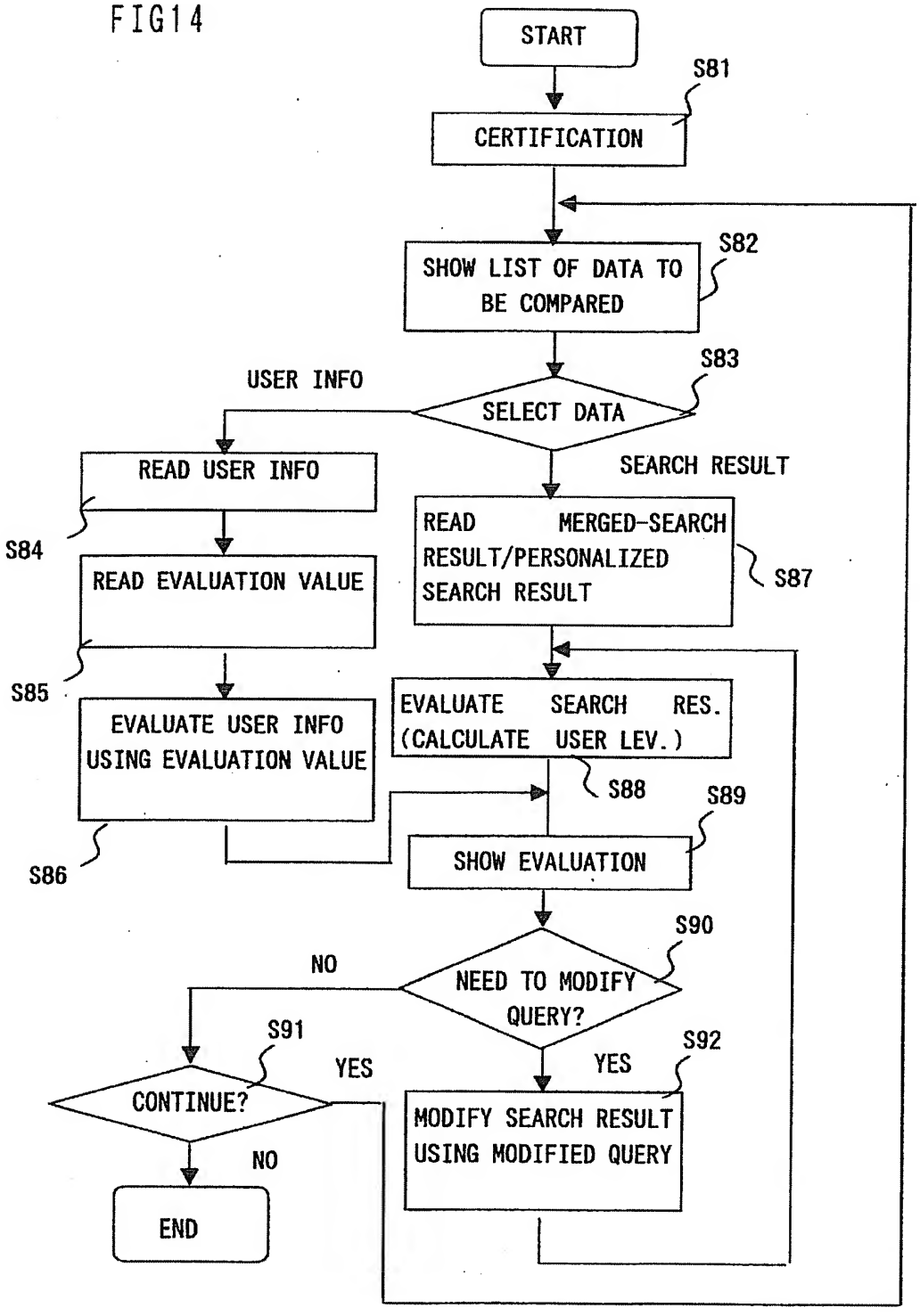

START
S81
CERTIFICATION
S82
SHOW LIST OF DATA TO BE COMPARED
S83
USER INFO
SELECT DATA
SEARCH RESULT
READ USER INFO
S84
READ EVALUATION VALUE
S85
EVALUATE USER INFO USING EVALUATION VALUE
S86
READ MERGED-SEARCH RESULT/PERSONALIZED SEARCH RESULT
S87
EVALUATE SEARCH RES. (CALCULATE USER LEV.)
S88
S89
SHOW EVALUATION
S90
NO
NEED TO MODIFY QUERY?
YES
S91
CONTINUE?
YES
NO
S92
MODIFY SEARCH RESULT USING MODIFIED QUERY
END

FIG15

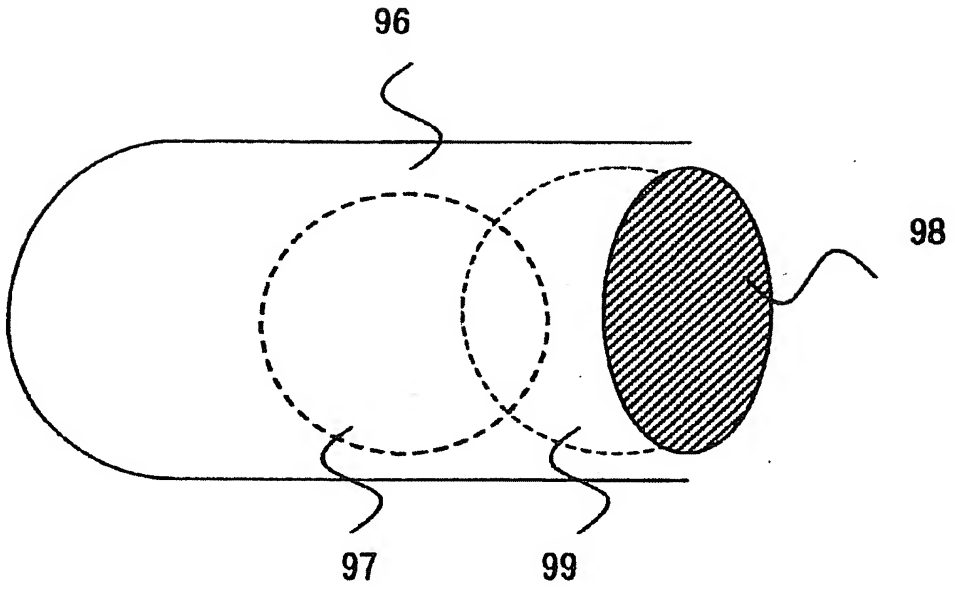
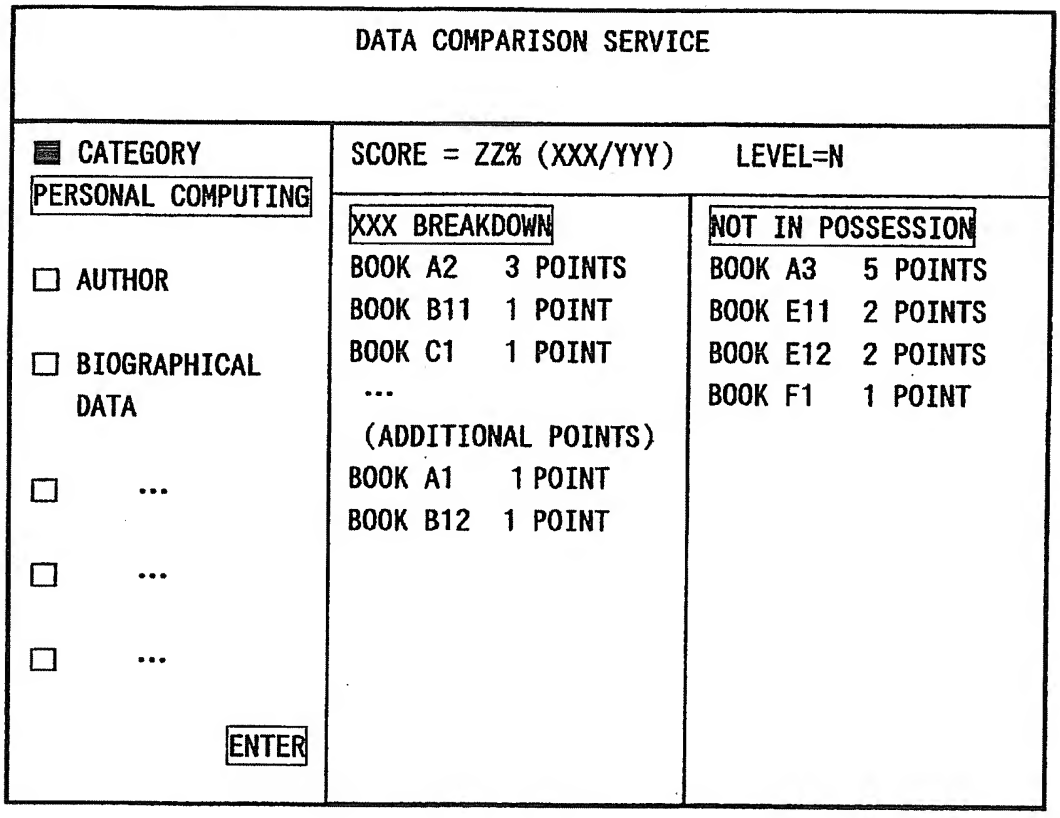


FIG16


DATA COMPARISON SERVICE
CATEGORY
PERSONAL COMPUTING
AUTHOR
BIOGRAPHICAL DATA
...
...
...
ENTER
SCORE = ZZ% (XXX/YYY) LEVEL=N
XXX BREAKDOWN
BOOK A2 3 POINTS
BOOK B11 1 POINT
BOOK C1 1 POINT
...
(ADDITIONAL POINTS)
BOOK A1 1 POINT
BOOK B12 1 POINT
NOT IN POSSESSION
BOOK A3 5 POINTS
BOOK E11 2 POINTS
BOOK E12 2 POINTS
BOOK F1 1 POINT

DATA SUPPLY CONTROLLING DEVICE, METHOD, AND STORAGE MEDIUM WHICH FACILITIES INFORMATION SEARCHING BY USER

FIELD OF INVENTION

The present invention relates to a data supply controlling device, a data supplying method, a storage medium storing a data supplying program, and a data supplying system, in which a system is used which searches for information (of books, for instance) using a data supplying device and provides search results to the user (user terminal) via the data supply controlling device in such a manner to provide information suitable for user's preference and interest by referring to user data and also to achieve improvement in both user privacy protection and communications efficiency.

BACKGROUND OF THE INVENTION

Recently, a system that enables a user to search for information (of books, for instance) on a network and place a necessary order (of books) is in practical use. In this type of system, first of all, the user sets various search conditions such as title, author's name, publisher, category, published year etc., by using a user terminal (personal computer, for instance). Then after a data search device searches for information on the basis of the search conditions, the user selects books he/she wants from the candidates searched for, by using the user terminal. And referring to the reviews etc., the user places an order. Distribution of books in electronic data form is growing today, since enormous number of titles has been published.

So, systems disclosed by Japanese Laid-Open Patent Application No. 11-259512/1999 (Tokukaihei 11-259512; published on Sep. 24, 1999) and Japanese Patent Publication No. 2976219/1999 (Tokkyo 2976219; published on Nov. 10, 1999) are examples of the system to provide information via the network as above.

In the system disclosed by Japanese Laid-Open Patent Application No. 11-259512, by a data search device, information is searched for while a query is concealed, and then unnecessary search results are deleted to make a list of the results shorter.

Nowadays, on account of the widespread use of the Internet, while a lot of information has been available from the net, security on the network has become an important issue. That is to say, although search conditions (query, for instance) are uniquely set by each user and are privacy-associated matters, the search conditions have been leaked to outsiders through line tracing etc. However, the system as above can ensure the confidentiality of the search conditions, since the concealment of the query makes the user's intention of the search blurred, and also the system enables the user to designate any of the search conditions as confidential.

In the system disclosed by Japanese Examined Patent Publication No. 2976219, the user registers one's user data such as personality, interest etc. in the host computer in advance, and on the basis of the user data, commercial information is searched for from data sent to the host computer every day by various information service agencies such as news organizations and data banks. Communications efficiency can be improved by using this system, because once the user data is registered, there is no need to input the data every time the user searches for commercial information. Also, since the search is done on the basis of the user data, the system can provide commercial information suitable for each user, and contributes to the user's convenience.

However, in the system disclosed by Japanese Laid-Open Patent Application No. 11-259512, even though the privacy of the user is protected on account of the concealment of the query, communications efficiency is low, because every time the user tries to search, the data necessary for the search has to be input from the user terminal. As a result, the user is forced to pay higher communications costs. This kind of problem becomes particularly noticeable in the search of books that requires large amount of data to be input.

On the other hand, in the system disclosed by Japanese Examined Patent Publication No. 2976219, while communications efficiency is improved since it is not required to input the user data more than once, the privacy protection is not sufficient, due to the lack of measures against the leakage of the registered data.

In other words, in the conventional systems as described above, the user cannot be assured of both the privacy protection and the efficient provision of useful information, when information is provided.

SUMMARY OF THE INVENTION

The present invention has an objective to offer a data supply controlling device, a method of supplying data, a storage medium storing a data supply program, and a data supply system, which achieve both enhanced user privacy protection and improved communications efficiency in supplying information (search result) which is suitable to user.

A data supply controlling device in accordance with the present invention, in order to accomplish the foregoing objective, is characterized in that it includes:

- a user data storage section for storing user data for each user who is to be supplied with information through a user terminal;
- a query generating section for generating, based on at least the user data, a query for a data supplying device to search for the information;
- a search result storage section for storing a result of a search conducted by the data supplying device for the information in accordance with the query;
- an identifying data management section for managing identifying data by which each user can be identified, separately from user specifying data by which each user can be specified; and
- a data supplying section for generating a search result matched with the identifying data of the user from the search result stored in the search result storage section for supply to the user terminal.

With the arrangement, the query generating section generates a query based on at least the user data stored in the user data storage section. Examples of the user data may include the age, gender, and areas of interest of the user, and books owned by the user. The data supplying device searches for information in accordance with the query, and the search result is stored in the search result storage section. Of the search result stored in the search result storage section, the data supplying section supplies only the search result matched with the identifying data of the user to the user terminal.

Since the query is generated in this manner based on user data, and the data supplying device searches for information based on the query, a search result that is suitable to each user can be presented to the user. Further, once stored in the

user data storage section, the user data of each user does not need to be input repeatedly. This improves communications efficiency between the user terminal and the data supply controlling device.

In addition, in the arrangement, the identifying data management section manages (for example, registers therein) identifying data, such as the user ID and certification number of the user, by which each user can be identified, separately from user specifying data, such as the name and address of the user, by which each user can be specified. This makes it difficult to specify individual users only by the identifying data even if the identifying data managed by the identifying data management section leaks out of the device or a third party infiltrates into the system, for example.

Consequently, with the arrangement, information suitable to each user can be presented to the user, and at the same time enhanced user privacy protection and improved communications efficiency can both be achieved.

A data supply system in accordance with the present invention, in order to accomplish the objective, is characterized in that it includes:

the data supply controlling device above; a data supplying device for conducting a search for the information in accordance with a query given by the data supply controlling device to provide a search result to the data supply controlling device; and

a user terminal for receiving a search result suitable to each user from the data supply controlling device,

the data supply controlling device, the data supplying device, and the user terminal being interconnected through a communications network.

In the arrangement, the data supply controlling device, the data supplying device, and the user terminal are interconnected through a communications network. Thus, a system can be offered which provides both enhanced privacy protection and improved communications efficiency.

A data supply controlling device in accordance with the present invention, in order to accomplish the objective, is characterized in that it includes:

user data storage section for storing user data for each user who is to be supplied with information through a user terminal;

a query generating section for generating, based on at least the user data, a query to search data supplied from a data supplying device;

a search section for searching for the information in accordance with the query;

a search result storage section for storing a result of a search conducted by the search section;

an identifying data management section for managing identifying data by which each user can be identified, separately from user specifying data by which each user can be specified; and

a data supplying section for generating a search result matched with the identifying data of the user from the search result stored in the search result storage section for supply to the user terminal.

In the arrangement, the query generating section generates a query based on at least the user data stored in the user data storage section. Examples of the user data may include the age, gender, and areas of interest of the user, and books owned by the user. The search section searches the data supplied from the data supplying device for the information in accordance with the query, and the search result is stored in the search result storage section. Of the search result stored in the search result storage section, the data supplying section supplies only the search result matched with the identifying data of the user to the user terminal.

Since the query is generated in this manner based on user data, and the information is searched for based on the query, a search result that is suitable to each user can be supplied to the user. Further, once stored in the user data storage section, the user data of each user does not need to be input repeatedly. This improves communications efficiency between the user terminal and the data supply controlling device.

In addition, in the arrangement, the identifying data management section manages (for example, registers therein) identifying data, such as the user ID and certification number o the user, by which each user can be identified, separately from user specifying data, such as the name and address of the user, by which each user can be specified. This makes it difficult to specify individual users only by the identifying data even if the identifying data managed by the identifying data management section leaks out of the device or a third party infiltrates into the system, for example.

Consequently, with the arrangement, information suitable to each user can be presented to the user, and at the same time enhanced user privacy protection and improved communications efficiency can both be achieved.

A data supply system in accordance with the present invention, in order to accomplish the objective, is characterized in that it includes:

the data supply controlling device above;

a data supplying device for supplying, to the data supply controlling device, data to be searched by the data supply controlling device; and

a user terminal for receiving a search result suitable to each user from the data supply controlling device,

the data supply controlling device, the data supplying device, and the user terminal being interconnected through a communications network.

In the arrangement, the data supply controlling device, the data supplying device, and the user terminal are interconnected through a communications network. Thus, a system can be offered which provides both enhanced privacy protection and improved communications efficiency.

A method of supplying data in accordance with the present invention, in order to accomplish the objective, is characterized in that it includes the steps of:

storing, in a user data storage section, user data for each user who is to be supplied with information through a user terminal;

generating, based on at least the user data, a query for a data supplying device to search for the information;

storing, in a search result storage section, a result of a search conducted by the data supplying device for the information in accordance with the query;

registering, in an identifying data management section, identifying data by which each user can be identified, separately from user specifying data by which each user can be specified; and

generating a search result matched with the identifying data of the user from the search result stored in the search result storage section for supply to the user terminal.

With the arrangement, a query is generated based on at least the user data stored in the user data storage section. Examples of the user data may include the age, gender, and areas of interest of the user, and books owned by the user.

The data supplying device searches for information in accordance with the query, and the search result is stored in the search result storage section. Of the search result stored in the search result storage section, only the search result matched with the identifying data of the user is supplied to the user terminal.

Since the query is generated in this manner based on user data, and the data supplying device searches for information based on the query, a search result that is suitable to each user can be presented to the user. Further, once stored in the user data storage section, the user data of each user does not need to be input repeatedly. This improves communications efficiency between the user terminal and the data supply controlling device implementing the foregoing process.

In addition, in the arrangement, the identifying data management section registers therein identifying data, such as the user ID and certification number of the user, by which each user can be identified, separately from user specifying data, such as the name and address of the user, by which each user can be specified. This makes it difficult to specify individual users only by the identifying data even if the identifying data registered in the identifying data management section leaks out of the device, or a third party infiltrates into the system, for example.

Consequently, with the arrangement, information suitable to each user can be presented to the user, and at the same time enhanced user privacy protection and improved communications efficiency can both be achieved.

A method of supplying data in accordance with the present invention, in order to accomplish the objective, is characterized in that it includes the steps of:

storing, in a user data storage section, user data for each user who is to be supplied with information through a user terminal;

generating, based on at least the user data, a query to search data supplied from a data supplying device;

searching for information in accordance with the query;

storing a search result in a search result storage section;

registering, in an identifying data management section, identifying data by which each user can be identified, separately from user specifying data by which each user can be specified; and

generating a search result matched with the identifying data of the user from the search result stored in the search result storage section for supply to the user terminal.

With the arrangement, a query is generated based on at least the user data stored in the user data storage section. Examples of the user data may include the age, gender, and areas of interest of the user, and books owned by the user. The data supplied from the data supplying device is searched for the information in accordance with the query, and the search result is stored in the search result storage section. Of the search result stored in the search result storage section, only the search result matched with the identifying data of the user is supplied to the user terminal.

Since a query is generated in this manner based on user data, and information is searched for based on the query, a search result that is suitable to each user can be presented to the user. Further, once stored in the user data storage section, the user data of each user does not need to be input repeatedly. This improves communications efficiency between the user terminal and the data supply controlling device implementing the foregoing process.

In addition, in the arrangement, the identifying data management section registers therein identifying data, such as the user ID and certification number of the user, by which each user can be identified, separately from user specifying data, such as the name and address of the user, by which each user can be specified. This makes it difficult to specify individual users only by the identifying data even if the identifying data registered in the identifying data management section leaks out of the device, or a third party infiltrates into the system, for example.

Consequently, with the arrangement, information suitable to each user can be presented to the user, and at the same time enhanced user privacy protection and improved communications efficiency can both be achieved.

A storage medium for storing a program implementing a data supply process in accordance with the present invention, in order to accomplish the objective, is characterized in that it stores a program executed by a computer to implement a process by means of the foregoing method of supplying data.

In the arrangement, the process by means of the method of supplying data can be implemented by a computer (for example, a data supply controlling device) executing the program.

For a fuller understanding of the nature and advantages of the invention, reference should be made to the ensuing detailed description taken in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** is a block diagram in accordance with an embodiment of the present invention, showing a general arrangement of a data supplying system.

FIG. **2** is an explanatory view showing an example of an arrangement of user data registered in a user data DB of a search service center of the data supplying system.

FIG. **3** is an explanatory view showing an example of an arrangement of a search rule registered in a search rule DB of the search service center **3**.

FIG. **4** is a flow chart describing an outline of a process of the data supplying system.

FIG. **5** is a flow chart giving a detailed description of the process of user registration shown in the flow chart of FIG. **4**.

FIG. **6** is a flow chart giving a detailed description of the process of user data registration shown in the flow chart of FIG. **4**.

FIG. **7** is a flow chart giving a detailed description of the process of query registration shown in the flow chart of FIG. **4**.

FIG. **8** is a flow chart giving a detailed description of the process to generate queries shown in the flow chart of FIG. **7**.

FIG. **9** is a flow chart showing a process of operation in query registration, after the generation of queries as above.

FIG. **10** is an explanatory view showing how data is exchanged between user terminals, a search service center, and a data server of the data supplying system.

FIG. **11** is an explanatory view showing an example of the exchanged data.

FIG. **12** is a flow chart giving a detailed description of the data demand process shown in the flow chart of FIG. **4**.

FIG. **13** is a block diagram in accordance with another embodiment of the present invention, showing an arrangement of a data supplying system.

FIG. **14** is a flow chart showing how the data supplying system is operated.

FIG. 15 is a general explanatory view showing how a search result is corrected on account of a correction of queries.

FIG. 16 is an explanatory view showing a displayed example of a result of comparison provided for a user terminal.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Embodiment 1

The following description will describe an embodiment of the present invention, referring to figures. By the way, although information of books is taken as an example of data provided for the user in this description, this invention is not confined to providing the information of books. This invention can be applied to provide general information of such as cars, watches and other commodities, for instance, by adding a database.

FIG. 1 shows a general arrangement of a data supplying system in accordance with this embodiment. This system contains a user terminal 1, a data server (data supplying device) 2, and a search service center (data supply controlling device) 3, and all of them are connected to each other through a network (communications network) 4 such as the Internet.

By the way, in this system, usually multiple user terminals 1 owned by the users and multiple data servers 2 (depending on types of information to provide) are connected through the network 4. However, to clarify the arrangement of the user terminal 1 and the data server 2, one user terminal 1 and one data server 2 are shown in the figure. Of course, a data supplying system including one user terminal 1 and one data server 2 may be arranged.

The user terminal 1 is used by the user being provided data by the present system. What can be used as the terminal 1 are, for instance, devices owned by the users and can be connected to the network 4 such as a PC (personal computer), a mobile information terminal and a mobile phone, and also a dedicated terminal capable of being used in the system.

The data server 2 is a server storing data provided for the user and also searching data appropriate to each user, according to a search request from the search server center 3. What can be used as the data servers 2 are, for instance, various websites on the network capable of providing data for the users such as on-line bookstores, websites containing information about selling books, and publisher's websites.

The search service center 3 stores the user data of each user, and also generates queries to be provided for the data server 2, by using the user data. A search result of the data server 2 is sent to the search service center 3, then from the search service center 3, the result is provided for the user terminal 1 of the user, according to the user's demand. This search service center 3 functions as a device to control a search of the data server 2, and a transmission of data, being an object of the search, from the data server 2 to the search service center 3 (this process will be described at the end of an embodiment 2).

Concrete arrangements of the user terminal 1, the data server 2 and the search service center 3 are successively described as follows.

The user terminal 1 includes user data registration means 11, query processing means 12, search result management means 13, a controller 14, a display 15, and certification means 16.

The user data registration means 11 registers each user's user data necessary to generate queries in the search service center 3, by following instructions from the controller 14. The user data registered in the user data registration means 11 is transferred to the search service center 3 afterwards. A detailed description of the user data is given later.

The query processing means 12 instructs to transfer the search results, to the user terminal 1, sent from the data server 2 to the search service center 3 by following an instruction of the controller 14. In the present system, a data search of the data server 2 is periodically done, for instance, once in a week. However, the search results are not transferred to the user terminal every time the search is done but transferred when the user demands the search result.

The search result management means 13 stores the search results being transferred from the search service center 3 to the user terminal 1.

The controller 14 is operated by the user to register the user data, to demand the transfer of search results and also to input a user ID, a certification number etc., and the controller 14 includes such as a keyboard and a mouse, for instance. The display 15 offers a GUI (Graphical User Interface) for various operations and displays search results, and includes a PC monitor, for instance.

The certification means 16 is used to access the search service center 3 and to confirm user IDs and a certification numbers input by the user by using the controller 14, when data is provided.

The data server 2 contains a database 21, data management means 22, search processing means 23 and a search rule generator 24.

The database 21 stores data provided for the user (data searched for), and in this embodiment it is information about books. Provided that multiple data servers 2 are connected to the network, each database 21 of data servers 2 stores specialized data (apart from books, about cars and watches, for instance) to which each data server 2 is dedicated.

The data management means 22 manages mainly a renewal of data being stored in the database 21. For instance, when new titles are published, the data management means 22 instructs the database 21 to store the data of the new titles.

The search processing means 23 actually searches the database 21, in accordance with the queries sent from the search service center 3.

In the process to generate queries in the search service center 3 as above, the search rule generator 24 provides multiple types of search rules to be stored in a search rule DB 34 (discussed later) of the search service center 3. What the search rule regulates is a rule by which on the basis of the registered user data, the search service center 3 generates queries in accordance with types of searched data.

If the data server 2 is a server searching information of books as in the present embodiment, as the search rule described above, the search rule generator 24 sets a rule that enables the search service center 3 to generate queries, which are capable of searching information of books, on the basis of the user data. To be more specific, examples of search rules are such as:

"If the user's possession data is international travel series, generate queries capable of searching guide books about international travel."

"If the user lives in Osaka and is interested in travel, generate queries capable of searching guide books of travel around Kansai region."

If the data server 2 is a server searching information of cars, as the search rule described above, the search rule generator 24 sets a rule that enables the search service center 3 to generate queries, which are capable of searching information of cars, on the basis of the user data. To be more specific, examples of search rules are such as:

"If the user is under 25, generate queries capable of searching cars which are popular among twentysomethings."

"If the user is interested in camping, generate queries capable of searching four wheel drive cars and wagons."

Since the search rule generator **24** sets search rules in accordance with types of data provided by the data server **2**, after the search rule generator **24** transfers the search rule to the search service center **3**, a query generator **35** (discussed later) of the search service center **3** can recognize which search rule is capable to generate the queries to instruct the appropriate data server **2** to search. On account of this, even if there are multiple data servers **2** corresponding to types of the provided data, in accordance with the provided search rule, the query generator **35** generates queries that instruct an appropriate data server **2** to search for data corresponding to the provided search rule, and as a result, data that the users need can certainly be searched.

Also, even if new titles are published, the search rule generator **24** can provide a search rule, generating queries capable of searching the new titles, for the search service center **3**. Thus new titles can be searched by using the queries as above.

Next, the search service center **3** is described. The search service center **3** includes user ID management means **31**, a user data DB **32**, user data management means **33**, the search rule DB **34**, the query generator **35**, a query DB **36**, query management means **37**, search means **38**, search result DB **39**, search result management means **40**, data supplying means **41**, and user specifying data management means **42**.

The user ID management means (identifying data management means) **31** manages user IDs and certification numbers which are used as identifying data to identify each user of the search service of the present system.

The user data DB (user data storage means) **32** registers and stores user data, which is;

registered from the user data registration means **11** of the user terminal **1**;

transferred to the search service center **3**; and

unique to each user for whom data is provided via the user terminal **1**.

In the present system, each user's user data is registered and stored to correspond to each one's user ID in the user data DB **32**, because multiple user terminals **1** owned by each user send user data. By the way, a detailed description of the user data is given later.

The user data management means (user data management means) **33** manages a registration of user data with the user data DB **32**, and also manages modification and update of user data in the user data DB **32**.

The search rule DB (search rule storage means) **34** stores search rules, which are sent from the search rule generator **24** of an appropriate data server **2**, in each data server **2**. That is to say, the search rule DB **34** stores search rules regulating the process that the query generator **35** generates queries matched with data to be searched. By the way, a data arrangement of search rules is described later.

In accordance with each user's user data and search rules stored in the search rule DB **34**, the query generator (query generating means) **35** generates queries capable of searching data that the users want, through the data server **2**. A search rule is renewed every time new rule is provided from the search rule generator **24** of the data server **2**, so thanks to the latest search rule, the query generator **35** can always generate appropriate queries.

To be more specific, as the queries described above, the query generator **35** generates the first query to include information that the user requires in the search results and the second query concealing the first query, and then the query generator **35** provides these two for the data server **2**.

The first query contains a base query, an additional query, and an exclusive query. The base query is a condition to include data matched with user data in the search results. The additional query is a condition to include data missed by the base query in the search results. The exclusive query is a condition to exclude each piece of poorly related data from the search results. The query generator **35** generates the queries which enables to search for data related to the users with certainty, and based on the queries, the data server **2** can search for data. On account of this, useful information is surely provided to the users.

Furthermore, the second query includes a dummy query to include data unnecessary for the user in the search results. Providing the dummy query in conjunction with the first query for the data server **2** causes that it becomes difficult to extract the search results matched with the first query from the search results derived from both the first query and the dummy query, since the dummy query conceals the first query. Therefore, even if search results are leaked out from the system, confidentiality of the user's search results can be surely protected, since it becomes extremely difficult to specify data that the user really wants.

Moreover, the query generator **35** generates a filtering query to delete the search results matched with the dummy query from the search results stored in the search result DB **39**. On account of this, while concealing the first query by the dummy query, unnecessary search results can be deleted. So only information that the user really wants can be provided for the user in the end.

The query DB **36** stores specified queries, matched with each the user ID, generated by the query generator **35** as above. In this embodiment, among the queries, the base query, the additional query and the dummy query are stored in the query DB **36**. Meanwhile, the exclusive query and the filtering query are registered in the user data DB **32**.

The query management means (query management means) **37** manages;

storing of queries in the query DE **36**; and

modification and update of queries stored in the query DB **36**,

merges and manages;

queries for each user generated by the query generator **35**, and provides;

merged query for the data server **2**.

The search means **38**:

keeps track of the data server **2** capable of doing a data search by using queries generated in accordance with a specified search rule;

provides queries specified by the query management means **37** for the data server **2**; and

gives instructions to the data server **2** to search data in accordance with the queries.

The search result DB (search result storage means) **39** exclusively manages and stores search results sent from the data server **2** that conducted searching by using the queries as above.

The search result management means (search result management means) **40** manages correlations:

between data (search result) stored in the search result DB **39** and identifying data of the user; and between the search result as above and the queries matched with each the user ID stored in the query DB **36**.

In this system, a data search is not done on a one-by-one basis by using each user's query but done integrally, by

virtue of queries of each user being merged in the query management means **37** and a merger query being provided for the data server **2**. Therefore, the search result DB **39** stores search results of all users.

In accordance with a requirement of search results from the user terminal **1**, the data supplying means (data supplying means) **41** extracts a search result (stored in the search result DB **39**) matched with an identifying data (user ID) managed by the search result management means **40**, and provides the result for the user terminal **1**. That is, from search results stored in the search result DB **39**, the data supplying means **41** provides the search results, which are not deleted by the filtering query, for the terminal **1**.

If, for instance, the data supplying means **41** provides search results, matched with each user, for the user terminal **1** only when the user demands the results by using the user terminal **1**, it becomes unnecessary to send the search results to the user terminal **1** when the user does not need. So, risks that search results are leaked out due to communications etc. can be minimized, and the privacy protection is further ensured.

By the way, this embodiment basically adopts so-called pull-type system that the data supplying means **41** provides data for the user terminal **1** in accordance with demand (push) from the users. However, the present system may be arranged as push-type system that the data supplying means **41** actively provides data (at regular frequencies, for instance) even without user's demand.

The user specifying data management means (user specifying data management means) **42** manages user specifying data which can specify each user. Also, the means **42** is used when recording/reading the user specifying data to/from, for instance, a removable storage medium, and the means **42** manages the user data in a condition that online access from outside is not available. The user specifying data as above includes user's name, address, IP address etc., as examples.

The user specifying data is used when making contact with users is required in the system administration. That is to say, when users are offered the data supplying service of the present system, owing to the user specifying data, the users can receive announcements such as:

- a deletion of the provided data, which has not seen for a long time, from the search result DB **39**; and
- a billing when a chargeable service is provided.

So, although the user specifying data does not really influence the process of data providing itself, it becomes essential when it is necessary to make contact with the users.

In the present embodiment, user IDs and certification numbers managed by the user ID management means **31** and user specifying data managed by the user specifying data management means **42** are collectively called user registration data. In this manner:

- user IDs and certification numbers; and
- user specifying data are managed independently in this system.

Now, according to an input process of the user terminal **1**, a detailed description of user data, stored in the user data DB **32** of the search service center **3**, is given. FIG. **2** shows an example of a data structure of user data **51** stored in the user data DB **32**. Incidentally, what will be described here is user data **51** to search for information of books.

The user data **51** is a specific example of user data stored in the user data DB **32**, and arranged to include multiple number of sets of user ID data **52** matched with each user's user ID. In other words, the user data **51** is managed on user-by-user basis, according to each set of user ID data **52**.

Each set of user ID data **52** includes user attribute data **53**, user preference data **54**, user possession data **55**, book purchase data **56**, search history data **57**, data supply history data **58**, search result filtering data **59**, search result evaluation data **60**, and user defined data **61**.

The user attribute data **53** is data attributed to user's personal circumstance, such as age, gender, occupation, living area etc. The user preference data **54** is data about each user's interest and preference. The user possession data **55** is about books that the user owns. The book purchase data **56** is data of books that the user plans or wants to buy in the future.

The search history data **57** is data about a history of searching conducted on the data server **2**. The data supply history data **58** is, about the data already provided for the user, and an evaluation and a comment originally attached to the data.

The search result filtering data **59** is data required when the query generator **35** generates a exclusive query and a filtering query (both described later) which are necessary to delete unnecessary data included in the search results. Provided that the search result filtering data **59** is registered as user data, in the end only useful search results can be provided for the user with no unnecessary data contained.

The search result evaluation data **60** is data showing user's evaluation for the past search results. The user defined data **61** is data that the user can designate in relation to a data search, for each of the following items such as:

- an each user's search rule to generate queries: an alteration of the generated search rule; and
- frequency to conduct the search.

The user data management means **33** encrypts user data of each user in accordance with corresponding user ID and certification number, and manages the data in accordance with each set of user ID. So the user data management means **33** can decrypt the crypted user data only when an appropriate certification number is presented in a certification procedure. On account of this, privacy of the user data can be surely protected, since it becomes difficult to crack the user data even when the user data is leaked out.

Next, a data arrangement of a search rule stored in the search rule DB **34** is described. FIG. **3** shows an example of an arrangement of a search rule **71** stored in the search rule DB **34**. By the way, the search rule **71** here is a search rule related to information of books.

The search rule **71** is a specific example of a search rule stored in the search rule DB **34**, basically indicating a rule to get (search for) information of books, rather than the information itself.

The search rule **71** includes rules corresponding individually to bibliographical data **72**, review data **73**, grouping data **74**, keyword data **75**, publisher data **76**, author data **77**, biographical data **78**, interest data **79**, occupation data **80**, region data **81**, mode data **82**, and tendency data **83**. These kinds of data may be provided:

- from the data server **2** which is a data provision source;
- by a third party;
- as originally created ones; and
- as a combination thereof above.

The bibliographical data **72** indicates a rule to get bibliographical data such as title, author, publisher, price, and published date, corresponding to a book code etc. that is used in a general book DB etc. The data **72** also indicates a rule to get data of a reference etc., when a particular book is designated to be searched.

The review data **73** indicates a rule to get data of an assessment of books such as a difficulty level, detailedness, and a critique of books, as an example. This review data **73**

is used for generating queries to provide data matched with what the user designates and determining ranking of search results.

The grouping data **74** groups books regardless of classification by category, and indicates a rule to get data matched with each group. This grouping data **74** is utilized to provide data related/similar to the user possession data **55**, and also information of books that are related to the data being associated with the data **55**.

The keyword data **75** is a database of rules to extract major keywords from title and contents of each book. This keyword data **75** is utilized to provide data related to the contents registered as keywords in the user defined data **61** in the FIG. **2**.

The publisher data **76** indicates a rule to obtain data about titles of books that are published by each publisher and major fields in which books are published. The author data **77** groups titles of books and major works by author, and sets a rule to obtain data corresponding to each group.

The biographical data **78** is a database of rules to get information of related books on a figure-by-figure basis, including not only titles of books but also contents such as historical figures and celebrities.

The interest data **79** indicates a rule to classify books in accordance with general areas of interests, and to get data matched with each area. The occupation data **80** indicates a rule to get information of books such as business books that can be provided for the user in relation to his/her occupation.

The region data **81** is a database of rules to get data related to local areas, for instance, a book that contains highly regional data and an area that is used as a scene of novel etc.

The mode data **82** is a database of rules to get bibliographical data of books related to data which is in fashion in each period, provided that arbitrary periods are designated.

The tendency data **83** indicates rules to obtain data that cannot be classified by using the data from the bibliographical data **72** to the mode data **82**, for instance, such as:

- information that actually many users demonstrate a propensity to set as the search condition etc., despite that the user base cannot be specified; or
- information that a specific user base often use, regardless of the difference of the category.

Next, an operation of the present system is described as below.

What is shown in FIG. **4** is the general operation of the present system. That is to say, first of all, the user selects a process to do in the present system by the user terminal **1** (step **1**; "step" will be referred to as "S"). On account of this, following the selection as above, either one process of user registration (S2), user data registration (S3), query registration (S4), and data provision request (S5) is done. After the process is done, the user selects whether to move to another process or not (S6). If the user continues after S6, the operation returns to S1, and if the user does not continue, the operation is terminated.

By the way, although the operation is usually done in order as S2, S3, S4, S5, any processes can be omitted, if not necessary. That is to say, for instance, provided that it is unnecessary to modify the registered data in either process of user registration (S2), user data registration (S3) and query registration (S4), it is possible to move from selection of the process (S1) to data provision request (S5) directly. Also, when the registered user data is required to be modified, it is possible to do the data provision request (S5) process after the user data registration (S3) process is done.

A detailed description of the processes presented above, user registration (S2), user data registration (S3), query registration (S4), and data provision request (S5), is successively given in the following paragraphs.

User Registration

FIG. **5** is a flow chart showing the procedure of user registration shown in S2 in FIG. **4**. First of all, the user selects whether it is new registration or not, by using the user terminal **1** (S11). If new registration is selected in S11, the user enters the aforementioned user specifying data for specifying each user (S12). Then a user ID is allocated to register the new user by the certification means **16** (S13), furthermore, a certification number in accordance with the user ID is set (S14). The user specifying data, the user ID and the certification number as above are sent to the search service center **3**.

Then the user specifying data management means **42** of the search service center **3** saves the user specifying data of each user in the storage medium (S15). Meanwhile, the user ID management means **31** registers the user ID and the certification number matched with the user specifying data (S16).

If new registration is not selected in S11, the user enters the user ID and the certification number (S17). On account of this, a certification procedure of the user is done by the certification means **16** (S18), and the user is confirmed as an authorized user (S19).

If the certification of the user is properly done in S19, the user specifying data management means **42** of the search service center **3** reads user specifying data stored in a specified storage medium (S20). Then if this data is required to be modified, responding to an operation of modifying in the user terminal **1**, the user specifying data management means **42** modifies the user specifying data (S21). Then the user specifying data management means **42** saves the modified user specifying data in the storage medium above again (S15). Also, the user ID and the certification number registered in the user ID management means **31** are registered in the user ID management means **31** again, as a user ID and a certification number matched with the modified user specifying data (S16).

If the certification of the user is not properly done in S19, the certification means **16** determines whether a number of the certification procedure being done is within retry limit or not (S22). If the certification procedure is done no more than the retry limit in S22, the certification procedure is repeated in S17. However, if the certification procedure is done more than the retry limit, the certification means **16** cancels the process as well as disconnects the system from the network **4** (S23). In this case, of course center administration means (not illustrated) of the search service center **3** checks a log of the access above and takes countermeasures against the unauthorized access.

By the way, although the description above is about the case that the user registration process is directly done online from the user terminal **1**, the user registration process may be done by the user or his/her proxy by using an alternative terminal, or it is also possible that a description of the registration sent from the user offline is registered in the search service center **3**.

User Data Registration Process

Next, the user data registration process is described. FIG. **6** is a flow chart showing the process of user data registration (S3) in FIG. **4**. By the way, the description will be given on the precondition that the user registration process shown in FIG. **5** is already done.

First of all, in the certification procedure using a user ID and a certification number of the user, the certification means **16** of the user terminal **1** confirms that the user is an

authorized user (S31). Incidentally, the procedure may be omitted, if this certification procedure is done somewhere from S2 to S5 in FIG. 4 and it is followed by the process of user data registration (S3). Also, if the certification procedure is done more than the retry limit, the whole process may be canceled, as described in FIG. 5. The description about the certification procedure as above can be applied in the processes described as below: query registration and data demand.

Now, if user data is newly registered (S32), the user generates one's user data (S33). The generated user data is registered in the user data registration means 11 of the user terminal 1, and after being sent to the search service center 3, the user data is stored in the user data DB 32 by the user data management means 33 (S34)

Meanwhile, if user data is not newly registered (if user data is already registered) (S32), the user data management means 33 of the search service center 3 reads the user data which is already registered in the user data DB 32 (S35), and if any modification or addition to the user data is made through the user terminal 1, modified/added user data is newly set (S36). The newly-set user data is registered in the user data DB 32 again, by the user data management means 33 (S34).

Query Registration Process

Next, the process of query registration is described. FIG. 7 is a flow chart showing the procedure of query registration (S4) in FIG. 4. By the way, the description is given on the precondition that the user data registration process shown in FIG. 6 is already done, since access to the user data DB 32 is necessary.

First of all, when the user terminal 1 confirms that the user is an authorized user by virtue of input of a user ID and a certificate number and also in the certification procedure (S41), the user selects whether generating new query or not (S42).

If the generation of new query is not selected in S42, the user terminal 1 reads already registered query matched with the user ID from the query DB 36 through the query management means 37 (S43). Then the user updates/modifies the read query as above by using the user terminal 1 (S44).

The query generator 35 of the search service center 3 reads data of modification/updating of the queries that are set by using the user terminal 1 and also data that registered in the user data DB 32, the search rule DB 34, and the query DB 36 that are needed to modify/update the queries (S45), and the query generator 35 automatically generates queries matched with the user ID (S46). By the way, a detailed description of the automatic generation of queries is given later.

Then the query management means 37 registers the queries generated in S46 in the query DB 36 again. Meanwhile, the user data management means 33 registers data, for providing data matched with a user ID (data to generate the exclusive query and the filtering query that are both described later), in the user data DB 32, as search result filtering data 59 (S47).

If the user selects to generate new query in S42, the user sets basic conditions to generate a query automatically, such as an update schedule of provided data (how often a data search is conducted), a maximum number of items of data to be provided, a search rule etc., by the user terminal 1 (S48). While the aforementioned data of the update schedule and the maximum number of items of data to be provided is stored in the user data DB 32 as the user defined data 61, for instance, the search rule above is stored in the search rule DB 34.

Then the query generator 35 reads data, registered in the user data DB 32 and the search rule DB 34, and also required to generate the query (S45). After this process, the following processes are the same ones as S46 and S47 as above.

Next, a detailed description of the aforementioned automatic generation of a query is given. By the way, what is described below is an automatic generation of a query which is already generated but is a subject to be modified/updated. As for an automatic generation of a newly-generated query, the procedure is same as the modification/update of queries except no query is read from the query DB in S45.

FIG. 8 gives a detailed description of the processes shown in FIG. 7: namely,

reading data from DBs in S45;

generation of the queries in S46; and

registration of data in DBs in S47.

Now, to make the description easier to be understood, the user data DB 32 at least stores data below as user data, in a provisional manner.

Gender: female

Age: 25

Occupation: fashion industry

Living Area: Osaka

Interest: travel

Owns: international travel series

Books Planning to Buy: novels by novels by novelist A

User Defined Data: except male fashion magazine

After data required to generate queries is read from the user data DB 32, the search rule DB 34, and the query DB 36 in S45 in FIG. 8, first of all, the query generator 35 generates a base query that effectively covers data provided for the user (S46-1).

In this process, the query generator 35 selects a search rule, capable of searching for data which is highly probable to be related to the user possessing the user data above, from the search rules that are read above. For this reason, the selected search rule should be able to search for an information magazine of Kansai region, a fashion magazine, travel guidebook etc., inferring from gender, age, occupation, living area and interest in the user data above. Then according to the selected search rule, the query generator 35 generates a query capable of searching for an information magazine of Kansai region, a fashion magazine, travel guidebook etc., as the base query. The base query above may be such as:

(Kansai+fashion+travel)×(information magazine)

Next, the query generator 35 generates an additional query to include data, which is highly probable to be excluded by the base query, in the queries (S46-2). For instance, as for the book planned to be bought "novels by novelist A", a search in accordance with the user data can be done without omissions, due to generation of an additional query capable of searching for the novels, by the query generator 35.

Then, the query generator 35 generates an exclusive query that excludes data which is poorly related to the base query and the additional query (S46-3).

If no supplementary information is provided, all information about fashion-related books is provided regardless of gender specificity, since the registered occupation is "fashion industry". So owing to the registration of the user defined data as "except male fashion magazine", the query generator 35 generates a query to exclude male fashion magazines from the search, as the exclusive query. The exclusive query above can be construct from keywords such as (female), for instance.

Also, as the exclusive query, the query generator **35** may generate a query to exclude books similar to "international travel series" that is registered as owns.

By the way, apart form the automatic generation by the query generator **35** alone, the exclusive query may be generated by a manual operation by the user.

Next, the query generator **35** selects any query that is already registered in the query DB **36**, and as a dummy query, the generator **35** adds the selected query to the queries above that are matched with the user ID (**S46-4**). So, data matched with the user ID in the user data DB **32** cannot be specified or inferred from the query, even if the query, matched with the user data registered in the query DB **36** and managed by the query management means **37** is leaked to outsiders.

To be more specific, according to the search rule DB **34** and data in the query DB **36**, a query capable of searching information, that is not related to an occupation, gender of the user and books owned by the user such as books on electricity, books on golf, novels by novelist B, is selected from queries registered in the query DB **36** by the query generator **35**, as the dummy query.

Finally, the query generator **35** generates a filtering query conducting a filtering process to exclude data, included by virtue of the dummy query generated in **S46-4**, from the search result (**S46-5**). That is to say, in the example above, as the filtering query, the query generator **35** generates a query which is able to exclude information such as books on electricity, books on golf, novels by novelist B from the search result.

Among these queries generated above:

- the exclusive query and the filtering query are registered in the user data DB **32** which is accessed after the certification by the user with reference to user ID; and
- the base query, the additional query and the dummy query are tentatively registered in the query DB **36**, after matching with the user ID and search conditions by the query management means **37**.

The exclusive query and the filtering query are both queries to delete data which is unnecessary when a search result is provided for the user, while the base query, the additional query and the dummy query are essential for a data search. Therefore, it is more manageable to divide queries into the combination of the exclusive query and the filtering query and the combination of the base query, the additional query and the dummy query such as the present embodiment, because of the difference of the functions.

Also, the latter group of queries being "tentatively" registered means that the queries are stored without merge at this moment. This system aims to increase communications efficiency on account of the merge of queries matched with each of the user terminals **1** and the delivery of a single query to the data server **2**. So it is inefficient to integrate the queries and store in the query DB **36**, every time the user modifies the registered data etc. Thus there is a step of the tentative registration in the present system so that queries are integrated just before being sent to the data server **2**, and in the end the integrated query is registered in the query DB **36**.

Now, more detailed description of the method of the generation of basic queries above is given. Assuming that the present system periodically (once in a week, for instance) conducts a data search in the data server **2** and the search result of the last time or the past search results is/are stored in the search result DB **39** as search results **1** to n (n is an integer that is not less than 2).

First, the query generator **35** reads data matched with a search rule registered in the search rule DB **34** from the user data DB **32**, and generates queries **1** to m (m is an integer that is not less than 2) that are candidates for the base query. After generating the queries **1** to m, the query generator **35** searches each of the search results **1** to n which are the past results stored in the search results DB **39**, by the queries **1** to m. Then the user terminal **1** gets feedback from the query generator **35** about the results of it.

Based on all of the results acquired above, the user briefly evaluates the effectiveness of each of the queries **1** to m. On account of this, each of the query **1** to m gets x (m, n), which is an evaluation value of each of the queries with respect to the search results **1** to n.

Incidentally, since it is too complicated for the user to evaluate the effectiveness of the queries on all of the search results **1** to n as done above, for instance, it is possible to extract some of the search results **1** to n at random and only evaluate the queries **1** to m with respect to the extracted search results. So to lighten the workload for the user, here search results **1** to l are extracted from the whole results **1** to n, and the queries **1** to m are evaluated with respect to them.

According to the evaluation value x (m, l) acquired from the evaluation of the queries **1** to m, the query generator **35** calculates evaluations E (**1**) to E (m) on each of the queries **1** to m. For instance, an evaluation E (m) is given by eq. 1,

$$E(m) = \frac{\sum_{k=1}^{l} x(m, k)}{l} \qquad \text{(Eq. 1)}$$

where the evaluation E (m) equals the sum of the evaluation value x (m, l) obtained by the evaluation of the query m on the search results **1** to l divided by the number of the search results **1** to l (l in this equation).

So, when the query generator **35** searches for data by adopting the queries **1** to m, to keep an adequacy of the number of sets of data provided for the user each time the data is updated, as the base query, the query generator **35** adopts a query matched with several of top-ranked evaluations among the evaluations E (**1**) to E (m).

In this manner, as the data provision is repeated, an effectiveness of the data provision by the base query is enhanced, on account of evaluations of actually provided data and periodical renewal of the base query, both conducted even after the data provision is started.

Next, based on a flow chart in FIG. **9**, what is described below is a process in the search service center **3** after the generation of queries as above. Assume that the process as below is supposed to be done periodically in the search service center **3**. This may be done simultaneously with the tentative registration of the queries, or done irregularly in accordance with a search schedule (not illustrated) stored in the query management means **37**.

First of all, the query management means **37** judges whether the queries (base query, additional query and dummy query) are tentatively registered in the query DB **36** in **S47** shown in FIGS. **7** and **8** or not (**S51**). If the queries are tentatively registered in **S51**, the query management means **37** analyzes the tentatively registered queries (**S52**). In other words, the query management means **37** extracts the queries (base query, additional query and dummy query) matched with a user ID from the query DB **36**. Then after extracting the exclusive query matched with the user ID from the user data DB **32**, the query management means **37** merges these four extracted queries (**S53**). By the way, merged query unifies queries when the queries share a keyword, or delete a query when it is included in another

query. Transition of queries on account of the merge will be described later.

By the way, the merged query inevitably includes a query that is irrelevant to a user ID, when the queries are merged. The aforementioned dummy query corresponding to this irrelevant query. Thus, the merge of four queries as above essentially equivalent to merging three queries (base query, additional query and exclusive query).

Then the query management means **37** stores the filtering query matched with each user ID, which is stored in the user data DB **32**, in a user ID table (not illustrated) included in the query management means **37** (S54). The filtering query above is a query to acquire search results matched with a user ID from the search results made by the merged query which is merged in S53. That is to say, the filtering query can exclude search results matched with the dummy query from the search result of the merged query.

So, the query management means **37** generates schedule data to conduct a data search by the queries registered in the query DB **36**, from:

- data about desirable time or duration of an update schedule (conduct of the search) designated by the user in the user defined data **61**; and
- data of an update schedule of the data server **2**, acquired from the search rules etc. provided by the data server **2**, then,

the schedule data is stored in a scheduling table (not illustrated) included in the query management means **37** (S55).

The query management means **37** registers queries which are not merged in S53 in the query DB **36**, and also modifies/deletes queries registered in the query DB **36**, when the query management means **37** determines that the queries are not used by a user possessing another user ID, by using a query management table (not illustrated) in the query management means **37** (S56).

Then the query management means **37** refers to the scheduling table in the query management means **37** (S57). So if a search is required, the search means **38** sends the merged query to the data server **2** and instructs the data server **2** the search. Consequently, following the instruction above, the search processing means **23** in the search server **2** searches the database **21** (S58). The data acquired from the search of the data server **2** is sent from the data server **2** to the search service center **3**, and registered in the search result DB **39** altogether (S59). Meanwhile, if the instruction to search is not required according to the search schedule, the process is terminated.

The search results registered in the search result DB **39** are not unique to individual users. So in accordance with a requirement of provision of search results from the user terminal **1** (described later), the data supplying means **41** extracts the search results matched with the user ID and managed by the search result management means **40**, from the search results stored in the search result DB **39**, and the extracted results are provided for the user terminal **1**. In this case, the data supplying means **41** can acquire the search results matched with a user ID of each user on condition that the means **41** applies the filtering query matched with the user ID to the search results stored in the search result DB **39**.

The present system does not provide a search result matched with each user until the user requires to provide the search result. So the search results in the data server **2** are stored in the search result DB **39** in a cohesive manner, and the registered search results are updated concurrently with the conduct of the search.

Meanwhile, if the queries are not tentatively registered in the query DB **36** in S51, i.e. the merged query are already registered in the query DB **36**, the operation is started from S57, and proceeded accordingly as above.

The following is a concrete description of the processes S54 and S56 as above. So, to make the description easier to be understood, it is assumed that the merged query contains, for instance, a query **1** including (female)×(fashion industry) and a query **2** including a sole keyword (X series by novelist A), and queries are updated weekly.

As for the query **1**, if a query related to the fashion industry is already included in the query DB **36**, the query management means **37** records the query number of the query above and the instruction to omit male fashion magazines in the user ID table. If there is no query related to the fashion industry in the query DB **36**, the query management means **37** stores the query **1** in the query DB **36** with no alteration, and also records the query number of the query **1** in the user ID table.

Like the query **1**, as for the query **2**, if a query about all novels by novelist A is already included in the query DB **36**, the query management means **37** records the query number of the query above and the instruction to restrict the search within X series in the user ID table. If no query related to novelist A is included in the query DB **36**, the query management means **37** stores the query **2** in the query DB **36** with no alteration, and also records the query number of the query **2** on the user ID table.

Furthermore, if the query DB **36** already contains any registered queries, the queries are merged with reference to a query possessing the highest update frequency, and queries registered in the query DB **36** are updated in S56.

In other words, if one user designates to conduct a search once a day and another user designates to do it once a month through designation of search frequency in the user defined data **61**, a merge of queries is conducted every day, in accordance with the search conducted once in a day.

By the way, if the data server **2** conducts a data search according to an instruction of the search in S58, the search result is registered in the search result DB **39** in accordance with the query number.

Now, transition of queries and search results in the present system is described, by using FIGS. **10** and **11**. Assuming that as the user terminals, there are a user terminal **1***a* in which a user ID corresponds to **001**, a user terminal **1***b* in which a user ID corresponds to **002**, and a user terminal **1***c* in which a user ID corresponds to **003**.

First of all, after user data **91***a*, **91***b* and **91***c* matched with each user ID are input from each of the user terminals **1***a*, **1***b* and **1***c*, the user data **91***a*, **91***b* and **91***c* are registered in the user data DB **32** of the search service center **3**, in accordance with each user ID. A description of the user data **91***a*, **91***b* and **91***c* is shown in FIG. **11**.

In the search service center **3**, the query generator **35** generates queries **92***a*, **92***b* and **92***c* for each of the users, in accordance with the user data **91***a*, **91***b* and **91***c*. Among the queries generated, a base query, an additional query and a dummy query are tentatively registered in the query DB **36**.

The query **92***a* contains the base query including (travel)× (overseas) and the additional query including (novelist A) only, that are generated in accordance with the user data **91***a* and a search rule in the search rule DB **34**. The query **92***b* contains a base query including (travel)×(industry), that is generated in accordance with the user data **91***b* and a search rule in the search rule DB **34**. The query **92***c* contains a base query including (travel)×(domestic), generated in accordance with the user data **91***c* and a search rule of the search rule DB **34**.

Then the query management means **37** merges these queries. In this example, the three base queries above can be unified to be a query **1** including (travel)×(international+industry+domestic), because the queries **92***a*, **92***b* and **92***c* share the keyword, (travel). As for the additional query in the query **92***a*, the query is left as a query **2**, since it cannot be merged because of no commonality between the queries **92***a*, **92***b* and **92***c*. A merged query **93** including the queries **1** and **2** is registered in the query DB **36**.

After the search means **38** sends the merged query **93** to the data server **2** at a certain timing and instructs a data search, the data server **2** conducts the data search in accordance with the merged query **93**. So a merged search result **94** is acquired, and the result **94** is sent to the search service center **3** and stored in the search result DB **39**.

When data provision is required from the user terminals **1***a*, **1***b* and **1***c* (described as below), the data supplying means **41** of the search service center **3** extracts a filtering query matched with each user ID from the user data DB **32**, and based on this query, search results matched with each user ID are acquired from the merged search result **94** stored in the search result DB **39**.

In this case, since the keywords included in the query **1** of the merged query **93** such as "domestic" and "industry" are not included in a base query and an additional query matched with the user ID **001**, queries such as (travel)×(domestic) and (travel)×(industry) are considered as dummy queries, with reference to the query **92***a*. Similarly, since the keywords included in the query **1** such as "overseas" and "domestic" are not included in a base query and an additional query matched with the user ID **002**, queries such as (travel)×(overseas) and (travel)×(domestic) are considered as dummy queries, with reference to the query **92***b*. Furthermore, since the keywords included in the query **1** such as "overseas" and "industry" are not included in a base query and an additional query matched with the user ID **003**, queries such as (travel)×(overseas) and (travel)×(industry) are considered as dummy queries, with reference to the query **92***c*.

In this manner, provided that a filtering query to exclude search results matched with each of the dummy queries is generated for each user ID and the filtering query matched with each user ID is applied to the merged search result **94**, data **95***a*, **95***b* and **95***c* matched with the respective user IDs are able to be acquired for supply to the users. The data **95***a*, **95***b* and **95***c* are provided for the user terminals **1***a*, **1***b* and **1***c* corresponding to the respective user IDs.

Data Provision Requirement

Next, what is described by using FIG. **12** is processes in which the user requires the data server **2** to provide data (search results) and the data is provided for the user.

First, the user terminal **1** is connected to the search service center **3**, and a certification procedure is conducted using a user ID and a certification number (**S61**). On account of this, the user is confirmed as an authorized user.

Then the search service center **3** presents a list of data provided for the user terminal **1** (**S62**). The provided data is newly arrived data, registration/history data, for instance. The newly arrived data is the latest search results in the data server **2**. The registration/history data is about user data or queries registered in the user data DB **32** or the query DB **36**, or about a search history.

The user selects type of data desired to be provided by using the user terminal **1** (**S63**). If the user selects to be provided with the newly arrived data through the registered queries in **S63**, the data supplying means **41** extracts data, matched with the user ID and managed by the search result management means **40**, from data (search results) registered in the search result DB **39** (**S64**) Meanwhile, if the user selects to be provided with the registration/history data in **S63**, the data supplying means **41** reads necessary data to be provided for the user from databases corresponding to sets of data (**S65**). After the data compiled or read in **S64** or **S65** is sent to the user terminal **1** by the data supplying means **41**, the data is displayed in detail on the display **15** of the user terminal **1** (**S66**).

Then the user selects a method to provide the data selected in **S63** (**S67**). As the method, in this case, a data browse and a data transfer are available. The data browse is a method to send data to the user terminal **1** promptly, so that the user can see the data immediately by using a browser of the user terminal **1**. The data transfer is a method to send the data as a file in which the data is saved altogether.

If the user selects to see data in the user terminal **1** in the manner of the data browse in **S67**, the data to be browsed is compiled (**S68**) and immediately transferred to the user terminal **1** (**S69**). On this occasion, in accordance with each user ID's filtering condition and data provision condition registered in the user data DB **32**, the data supplying means **41** of the search service center **3** narrows down and prioritizes the data to be provided, and as the data to be browsed, the results are provided for the user terminal **1**.

In the case that search results or registered data are required to be browsed immediately, or in the case that user data or queries are required to be modified/deleted immediately in accordance with the results of browsing of search results and registered data, provision of data by way of the data browse as above may be preferable.

Meanwhile, if the user selects the data transfer in **S67**, the data to be transferred is compiled (**S70**) and transferred to a certain user terminal **1** (**S71**). On this occasion, like the generation of the data to be browsed, in accordance with each user ID's filtering condition and data provision condition registered in the user data DB **32**, the data supplying means **41** of the search service center **3** narrows down and prioritizes the data to be provided, then encrypts the generated data and temporarily stores it, and finally transfers the data to be transferred to a specified terminal at a specified time. In the case that the provided data is used in another terminal or is printed out, provision of the data by way of the data transfer may be preferable.

Subsequently, if the user selects to continue the operation, the processes as above are repeated from **S62** (**S72**). If not selected, the operation is terminated.

By the way, as for communications means used in the data browse and the data transfer, using encryption of protocols such as SSL, the Secure Socket Layer, due consideration not to be intercepted by others is given of course.

As the description above shows, in the present embodiment, search results suitable for each user can be provided for each user, since queries are generated in accordance with user data of the user, and based on these queries, the data server **2** searches for data.

It is preferable that information of books that the user owns is included in the system to provide data by using a data search, especially when providing data related to books. However, information of books that the user owns can be regarded as highly private, thus it is desirable that the user manages the data by him/herself and utilizes for the data search. On the other hand, however, setting such data and exchanging it over the network each time the user searches data in multiple data providers cause considerable communications insufficiency and also force the user to pay a lot of communications costs, because of an immense amount of

the data and a requirement to set many elements properly to search for a desired book.

So, in the present embodiment, communications between the user terminal **1** and the search service center **3** can become efficient and communications costs of the user can be reduced, since once user data of each user is stored in the user data DB **32**, it becomes unnecessary to input the same user data again.

Also, the present invention is characterized in that (a) identifying data such as a user ID and a certification number and (b) user specifying data are managed separately. A probability of occurrence of leakage of the user specifying data to the outside is extremely low, because the user specifying data is stored in a storage medium being independent from the user ID and the certification number, and also is managed in a state that outsiders cannot access by the network **4**. Also, even if the identifying data registered in the user ID management means **31** is leaked out, it is virtually impossible to specify the user without the above user specifying data. Hence this system managing the user ID, the certification number, and the user specifying data separately can certainly protect each user's privacy, and the user can use the data provision service with a feeling of security.

In short, the present invention can provide data suitable for each user with both improved communications efficiency and privacy protection of the user.

Also, the present invention is characterized in that queries matched with each user are merged and provided for the data server **2**. Provision of the queries for the data server **2** requires a single transmission from a data supply controlling device. Thus, compared to an arrangement that each of queries matched with each user is provided for the data server **2** in separate transmissions, the present embodiment further improves communications efficiency and also reduce costs for users.

Moreover, if the query matched with the user is provided for the data server **2** in separate transmissions, the data server **2** has to conduct a search each time the server **2** receives a query. However, in the present embodiment, the server **2** only conducts a search when receiving the merged query. Hence a search in the data server **2** can be done efficiently.

Moreover, the user do not have to be accustomed to using the system etc., since:

- data required for the data provision is registered in the search service center **3**;
- appropriate queries are generated automatically; and the results of the processes above are provided for the user.

Also, a search can be conducted efficiently, because another query which is registered as a search rule or a query registered by another user is diverted as a dummy query which is generated as a query.

Furthermore, when the data server **2** is instructed to conduct searches for multiple users, the data server **2** and the network can be utilized efficiently, because the search service center **3** provides a merged query and then instructs the search altogether.

Embodiment 2

The following description will describe an another embodiment of the present invention, referring to figures. By the way, arrangements identical to those in Embodiment 1 that have the same numbers and the descriptions are omitted.

FIG. **13** shows a general arrangement of a data supplying system in accordance with this embodiment. As the figure shows, the arrangement of this embodiment is identical to that of Embodiment 1, except a data comparator **43** is provided in a search service center **3**. Following is a description of processes in this embodiment, based on a flow chart in FIG. **14**.

After a user terminal **1** is connected to the search service center **3** via a network **4**, if a process of data comparison is selected, first of all, the user is certified as an authorized user by the user terminal **1**, by a certification procedure using a user ID and a certification number (S81). Then the data comparator **43** reads a list of data that can be compared to (will be referred to as data to be compared) from a search rule DB **34**, and the data is shown in the user terminal **1** (S82).

Examples of the data to be compared are such as user data and a search result. Comparing user data means that, for instance, about books by a novelist, user possession data registered in a user data DB **32** and the number of currently published titles of the novelist are compared. Incidentally, a category (for instance literature, science, SF, foreign books, cartoon etc.) can substitute for the novelist.

Meanwhile, comparing search results means to compare the result searched for a merged query to the result matched with a user ID.

When the data to be compared is displayed in S82, the user selects data which the user wants to compare from the displayed data to be compared, and if another parameter such as a novelist or a category is required to be designated, the user designates it (S83).

If, for instance, a comparison of user data related to user possession data is selected and moreover a comparison of user possession data in relation to novelist A is designated in S82, the data comparator **43** reads novelist A-related user possession data of the user who instructs the data comparison, which is stored in the user data DB **32**, via user data management means **33** (S84)

Then the data comparator **43** reads data to be compared to the user possession data read in S84, i.e. an evaluation value (described later) for comparing the user possession data to the number of currently published titles of novelist A, from the search rule DB **34** (S85). Then the data comparator **43** evaluates the user data, by comparing the user data of the user (user possession data) to the evaluation value (S86). By the way, the method to evaluate user data is described later.

If a comparison of search results is selected in S83, the data comparator **43** reads a merged-search result produced by a merged query which is stored in a search result DB **39**, and also reads:

- a search result that is produced by the merged-search result above and will be provided for the user who instructs the data comparison; or
- search results that have been provided for the user (a history of the search results) (S87).

Then the data comparator **43** compares the merged-search result to the search result matched with each user, and evaluates the search result that will be provided for the user or the search results that have been provided for the user, i.e. calculates a level of the search result (S88).

A level of the search result is, for instance, a total sum of evaluation values to which each set of data is assigned weight, considering an amount of data in a specified category and also generality, detailedness and difficulty level of a content of each set of data. So in accordance with this total sum, the evaluation values are represented as, for instance, an amateur level, a standard level and an expert level etc.

Especially, data of the type and field etc. of book such as magazines, books in general and specialized books are already encoded and managed in the case of books, so on the basis of this, it is possible to classify books. It is also possible to estimate generality, detailedness and difficulty of the contents of a book from the number of pages and price etc. Furthermore, data etc. in relation to a target book may

be provided. This kind of data can be provided from the data server **2** to the search service center **3**.

For instance, if book data classified in a category of personal computer, which is provided for the user, is restricted to book information of magazines and introductory books by an exclusive query etc., the level of the search result is represented as an amateur level. If the data provided for the user above contains explanatory books on an OS or software, the level of the search result is represented as a standard level, and if it contains information of specialized books about programming theories, the level of the search result is represented as an expert level. By the way, the search results that have been provided for the user are evaluated in the same manner.

Then data supplying means **41** registers the result of the evaluation that is done by the data comparator **43** in **S86** or **S88**, in the user data DB **32**, and also provides the result of the evaluation for the user terminal **1**, in accordance with a data requirement of the user (**S89**). In this case, the result of the evaluation can be provided for the user terminal **1** in the same manner as that of the data provision requirement described in Embodiment 1 using the flow chart in FIG. **12**. Then taking the provided result of the evaluation into account, the user selects whether to modify a query or not, as needs come up (**S90**).

At this point, if the target to be compared to is user data, the operation is moved to **S91**, since the modification of a query does not affect an evaluation of user data. The user selects whether the data comparison process is continued or not in **S91**. If the process is not continued, the operation is terminated, and if the process is continued, the operation is returned to **S82** to conduct the same processes.

Meanwhile, if data to be compared is a search result, the user may need to reassess the query to get an appropriate search result again, depending on the evaluation shown in **S89**. In this case the modification of the query may be selected by the user in **S91**. In this case, for instance, a query generator **35** modifies an exclusive query, and data supplying means **41** modifies the search result provided for the user (**S92**).

That is to say, as shown in FIG. **15**, not-yet-modified search results **98** which are seen as in relatively low level poorly represent the whole merged search results **96**, so it can be considered that the results **98** have no/few overlapping result/results with standard results **97**. So, if the query generator **35** corrects, for instance, the exclusive query by excluding a keyword, a range of search results excluded from the merged search results **96** becomes narrower. On account of this, a post-correction search results **99** can cover wider range of search results than the not-yet-modified search results **98**, and also the results **99** can be similar to or overlap with the standard results **97**, slightly at least.

By the way, which keyword is excluded from the exclusive query may be designated, for instance, by the user, or by the data comparator **43** to get a search result at higher level, taking detailedness, a difficulty level, and generality as data into account.

After the modification of the search result in s **92**, the evaluation can be confirmed by evaluating the search results again in **S88**.

As described above, by comparing search results as above, it becomes possible to evaluate that whether the search result provided for each user is appropriate for a specified set of data (novelist, category) or not, and also possible to modify the result in accordance with the evaluation.

Also, if a level of search results is an amateur level, the query generator **35** can modify an exclusive query to make the level of the search results to be, for instance, a standard level. If the level of search results is a standard level, the query generator **35** can modify a query to make the level of the search results, for instance, to be an expert level. In this manner, the user can certainly obtain search results that are at higher level than the not-yet-modified search results, owing to the modification of a query (exclusive query, for instance) in accordance with the level of search results.

Now, a method to evaluate user data is described.

First of all, a general description of the method is as follows:

- the evaluation value of each book judged by the contents etc. is registered in the search rule DB **34**;
- a total point of each particular novelist or category is calculated in accordance with past search actions and the point is recorded; and
- to what extent the user has books in contradistinction to books and publications that are categorized in terms of already registered specific novelists and categories is evaluated, with the aid of the total point.

The following is more specific description of this.

TCa is given by eq. 2,

$$TCa = \frac{Pa}{TPa} = \frac{\sum_{l=1}^{m} px(a, l)}{\sum_{k=1}^{n} ax(a, k)}$$

where a basic result of comparison TCa referring to item a (specific novelist or category, for instance) can be represented as a ratio of TPa and Pa, Pa being a sum of evaluation values px (a, l) that are assigned to m books owned by the user and belonging to a, and TPa being a sum of evaluation values ax (a, k) assigned to all n books belonging to a in accordance with importance and levels of the books.

By the way, k is positive integers 1 to n, and l is positive integers 1 to m.

Incidentally, although the evaluation values ax (a, k) and the evaluation values px (a, l) are treated as different variables, these two are identical for the same book.

Also, for instance, an evaluation value for a representative work of a novelist may be different from those of other works as weights are assigned, since of course possessing the representative work may be valued higher than possessing other works.

For instance, in the case of novel etc., the degree of reputation of a novel can be estimated from data such as received awards and reviews on the book, and the number of published copies. These kinds of data can be provided from the data server **2** to the search service center **3**.

By the way, there are some books on technology (on electric circuit, for instance) and textbooks that can be seen as possessing equivalent level or contents. In this kind of case, if the user has a certain book on technology, the user may be seen as possessing other books on technology as well, even without actually other books that are in the same group as and having similar contents to the book above.

So, if some books can be seen as having equivalent contents or level, the equivalent books are registered in a single group, and if the user has any one book in the group, evaluation values px assigned to other books registered in the group are added to a sum of evaluation values Pa. On account of this, a knowledge level of the user in the field to which the book belongs can be judged by just whether possessing a book belonging to the group or not, regardless of the number of books in one's personal library.

Also, books related each other in terms of detailedness or a difficulty level can be treated similarly as above. In other words, if the user has a book on technology with a higher difficulty level, the user can be regarded as possessing a book on technology with a lower difficulty level as well, i.e. regarded as understanding the contents of the book, even if the user does not own the book with lower difficulty level.

Therefore, provided that:

- a correlation between multiple number of books in terms of detailedness and difficulty levels of the contents is registered; and
- a user owns highly detailed and/or difficult books, by adding evaluation values (px) of books possessing lower level than that of the books above (books subordinate to the books above) to a sum, Pa, of evaluation values, a knowledge level of the user in the field to which the book belongs can be judged regardless of the number of books in one's personal library, as is the case above.

The aforementioned equivalence and correlation between books can be determined through a procedure below, for instance. That is to say, based on statistics about the situation of data use etc. by each of the users who are equally ranked in terms of a specified category, it is possible to determine the equivalence and the correlation between books in consideration of a difference of frequency of the data use between the users and an inclusion relationship by data of the users at higher levels.

Also, accuracy of the data can be improved by carrying out surveys such as a questionnaire and voting.

These sorts of the statistical analysis and the surveys can be conducted in the data server **2** and the acquired data can be provided from the data server **2** to the search service center **3**. Also, the analysis and the surveys can be conducted in the search service center **3**. By the way, to improve accuracy of the data acquired from the surveys such as the questionnaire and voting, for instance, a person, looking after either the data server **2** or the search service center **3** that actually conducts the survey, can be involved with the planning of contents of the surveys.

Now, FIG. **16** shows a sample image displayed on a display **15** of the user terminal **1**.

First, in a frame on the left side of the screen, the user selects what is compared, designates details of the contents in a box, and starts the process of comparison by clicking "enter". The figure shows a case that "category" was selected to be compared, and the user typed the word "personal computing" as the detail, to compare sets of data related to the category.

At the top right of the screen, a result of comparison that is worked out by a process of data comparison of the present embodiment is displayed in a form of a score (zz% in the figure). This score is indicated as a ratio of:

- an evaluation score xxx that is a sum of evaluation values of one's personal library related to "personal computing"; to
- an evaluation score yyy that is a sum of evaluation values of all books related to "personal computing".

Also, a knowledge level of the user's books derived from contents and correlations of his/her books as described above is shown as level N. In this example, the knowledge level is separated into grades, for instance, H (high), N (normal) and L (low), based on the score above. The knowledge level in this example is normal.

What are displayed in the left side of an area immediately below the score are:

- a breakdown of the evaluation score xxx in relation to books owned by the user; and
- a list of books that are not owned by the user but the scores are added to, since the contents are similar or equivalent to the books that the user owns.

The example shows that although the user does not own books **A1** and **B12**, evaluation scores of the books **A1** and **B12** are added to, since the user owns books **A2** and **B11** in correlation with the books above.

Moreover, in the right side of the area immediately below the score, a breakdown of not-yet-owned books among books related to "personal computing" is listed in descending order of evaluation scores. The division of a breakdown of information of books as supplementary information enables the user immediately to acquire information about books that the user is required to read in the future, and also enables the user to realize more efficient reading. In addition, providing such supplementary information improves user's knowledge level in the selected field.

In this manner, in the present embodiment, each evaluation value of data, which is searched for in accordance with user data stored in the user data DB **32**, is stored in the search rule DB (evaluation value recording means) **34**, then based on the sum of the evaluation values above and the sum of evaluation values matched with the user data above, the data comparator (user data evaluation means) **43** evaluates the user data. On account of this, it becomes possible to judge a knowledge level of the user who is matched with the user data, in the data that is a target to be compared.

Moreover, the aforementioned data comparator (search result comparing means) **43** compares search results stored in the search result DB **39** and a search result matched with each user, and judges a level of the search result. This enables to judge effectiveness of a search result provided for each user.

By the way, the operations described in the embodiments above can be realized by a computer program. The program is stored in a computer-readable storage medium. In the present invention, this storage medium may be a memory (not illustrated), for instance ROM itself, that is required when the search service center **3** operates, or may be a program storage medium that is readable by inserting into a program reader provided as an external storage device (not illustrated).

In both cases, the stored program may be arranged to be executed by an access of a microprocessor (not illustrated), or arranged so that the program is executed by reading the stored program and then downloading the read program to a program storage area (not illustrated) of a delivery server and a receiving server. In this case, a program for downloading is stored in the main device in advance.

The program storage medium may be a storage medium that can be separated from the main device, and may be a medium that stores a program permanently, including tapes (magnetic tape, cassette tape etc.), magnetic discs (floppy disk, hard disk etc.), optical disks (CD-ROM, MO, MD, DVD etc.), cards (IC card (including memory card), optical card etc.), and semiconductor memories (mask ROM, EPROM, EEPROM, flash ROM etc.).

Also, the program storage medium may be a medium that stores a program temporarily, for a purpose of downloading the program from a communications network, since the system of the present embodiment is arranged to be connectable to communications networks including the Internet. By the way, in the case of downloading the program from a communications network, the program for downloading may be stored in the main device in advance, or may be installed from another storage medium.

Contents stored in the storage medium are not limited to a program but may be data as well.

Although the description above described the system in which search processing means **23** is provided in the data server **2** and the data server **2** is equipped with a function of an information search, search means **38** of the search service center **3** may have the function of the search processing means **23**, i.e. the function of an information search. In this case, data matched with a type of information provided for the user, stored in a database **21** of the data server **2** is sent to the search service center **3**, and based on a query generated by the query generator **35**, an information search is conducted by the search means **38**. The same effects as that of the present embodiment can be acquired from this embodiment.

Incidentally, Japanese Examined Patent Publication No. 2976219 discloses that user data is recorded, for instance, in a storage medium such as IC memory, and the user data stored in the storage media above is read by a terminal of the user when a search for commercial information is conducted, and then a host computer conducts an information search by picking up only required data from the user data. However, the invention above and the present invention are different in that while the invention above stores user data in a storage medium, the present invention only stores user specifying data (data which can specify each user, such as name and address) in a storage medium.

In the present invention, user data only matched with identifying data (correlation between identifying data and a removable storage medium is recorded only in the storage medium above) can be perpetually utilized in the search service center **3**. Also, queries are not generated every time an information search is conducted but queries registered in databases in advance are used. On account of this, when information is provided for the user, a connecting time for information provision via communications means is reduced, since it becomes unnecessary to take time to generate queries and total required time to search for information is reduced. However, a time to extract a search result from a merged query is needed.

A data supply controlling means in accordance with the present invention can be arranged to further include user specifying data management means for managing the user specifying data in such a manner that no access can be made to the user specifying data over a communications network.

According to the arrangement above, for instance, being stored in a storage medium that is removable from the main device such as IC memory, the user specifying data is managed by the user specifying data management means in such a manner that no access can be made to the user specifying data over a communications network. On account of this, a leakage of the user specifying data in the process of communications can be certainly prevented.

The data supply controlling means in accordance with the present invention can be arranged to further include query management means for merging and managing a query for each user generated by the query generating means, so as to supply a merged query to a data supplying device.

According to the arrangement above, queries matched with the users generated by the query generating means are merged by the query management means. For instance, if queries include the same keyword, the queries are merged into one query. In this manner, the merged query is provided for the data supplying device by the query management means. Hence the data supplying device uses the merged query for searching information.

In this manner, the provision of a query to the data supplying device requires a single communications from the data supply controlling device. So communications efficiency is improved, comparing to an arrangement that non-merged queries matched with each user are provided for the data supplying device in separate transmissions.

The data supply controlling device in accordance with the present invention can be arranged to further include user data management means for encrypting and managing the user data in relation to matched identifying data.

According to the arrangement above, user data becomes certainly protected and impossible to be decrypted even if the user data is leaked out of the device, since the user data management means encrypts and manages the user data in accordance with matched identifying data.

The data supply controlling device in accordance with the present invention can be arranged so that the query generating means generates, as the query, a first query to include information necessary to the user in a search result and a second query to conceal the first query, both for supply to the data supplying device.

According to the arrangement above, it is difficult to specify information that the user really requires, even if the search result produced by the data supplying device in accordance with the first and second queries is leaked out of the device, since the second query conceals the first query that provides required information for the user. Hence privacy of the user in relation to the search result can be protected.

The data supply controlling device in accordance with the present invention can be arranged so that the first query is composed of a base query to include information matched with the user data in a search result, an additional query to include information excluded by the base query in an information search in the search result, and an exclusive query to exclude each piece of poorly related information from the search result.

According to the arrangement above, information that the user requires can be provided with certainty by generating a query that can certainly search for information in relation to the user thanks to the inclusion of the base query, the additional query, and the exclusive query in the first query.

The data supply controlling device in accordance with the present invention can be arranged so that the second query is a dummy query to include information unnecessary to the user in the search result.

According to the arrangement above, the search results can include results that are poorly related to the user because the search by the data supplying device is conducted in accordance with the first query and the second query which is a dummy query that are sent to the data supplying device altogether. Hence even if the search result is leaked out of the device, privacy of the user in relation to the search result can be protected with certainty, since it becomes extremely difficult to specify information that the user really needs.

The data supply controlling device in accordance with the present invention can be arranged so that:

- the query generating means generates a filtering query to filter out a search result derived from the dummy query from search results stored in the search result storage means; and
- the data supplying means supplies the search result not to be filtered out to the user terminal from the search result storage means.

According to the arrangement above, unnecessary search results matched with the dummy query are filtered out by the filtering query, then the consequence is provided for the user terminal. On account of this, while concealing the first query by the dummy query, only information that the user really requires can be supplied for the user in the end.

The data supply controlling device in accordance with the present invention can be arranged so that the data supplying means supplies a search result suitable to each user to the user terminal in response to a data supply request sent from the user through the user terminal.

According to the arrangement above, the data supplying means can supply a search result suitable to each user to the user terminal only when, for instance, a data supply request is sent from the user through the user terminal. So privacy protection can be further accomplished, since it becomes unnecessary to send a search result to the user terminal when the user does not require so that an opportunity of a leakage of a search result in the process of communications etc. is minimized.

The data supply controlling device in accordance with the present invention can be arranged to include search rule storage means for storing a search rule specifying a rule by which the query generating means generates a query in accordance with kinds of information to be searched for, wherein:

the query generating means generates a query based on the user data and the search rule.

According to the arrangement above, the query generating means generates a query matched with types of information to be searched for, based on a search rule stored in the search rule storage means in addition to the user data. So this enables the data supplying device appropriately to search for information in accordance with types of information to be searched for in relation to the query and information matched with user data. As a result, information (search result) that the user really need can be acquired with certainty.

The data supply controlling device in accordance with the present invention can be arranged so that the search rule stored in the search rule storage means is supplied from the data supplying device and is updated upon every supply.

According to the arrangement above, the search rule supplied from the data supplying device is stored in the search rule storage means. Also, the search rule stored in the search rule storage means is updated upon every supply of the search rule from the data supplying device. On account of this, the query generating means can recognize which search rule should be selected to generate a query that makes an appropriate data supplying device search for information. Hence even if there are more than one data supplying device in accordance with types of information to be provided, an information search matched with the search rule can certainly be done in a specified data supplying device that provides the search rule.

The data supply controlling device in accordance with the present invention can be arranged to include:

evaluation value storage means for storing an evaluation value for each piece of information to be searched for in accordance with the user data; and

user data evaluation means for evaluating the user data based on a sum of all the evaluation values and a sum of the evaluation values matched with the user data.

Provided that books by novelist A owned by the user is the user data and all books by novelist A registered in data supplying device is data to be searched for in accordance with the user data, the evaluation value storage means stores an evaluation value for each piece of information of all books by novelist A. Then the user data evaluation means evaluates the user data based on a sum of the evaluation values of those books and a sum of the evaluation values matched with the books owned by the user. On account of this, a knowledge level of the user matched with the user data, about the information to be compared (in this case novelist A) can be judged.

The data supply controlling device in accordance with the present invention can be arranged to include search result comparing means for comparing the search result stored in the search result storage means with a search result suitable to the user, so as to determine a level of the search result suitable to the user.

According to the arrangement above, compared to the search result stored in the search result storage means, a level of the search result suitable to the user that is stored in the search result storage means can be judged by the search result comparing means. Examples of the level of the search result are such as an amateur level, a standard level and an expert level, and the level is comprehensively determined in consideration of, for instance, detailedness, a level of difficulty, and generality of the search result with a certain degree of weights assigned. So referring to the level of the search result, the user can easily judge whether a search in a higher level is required to be done or not, i.e. whether the query should be reassessed or not. In other words, this enables to judge effectiveness of the search result provided for each user easily.

The data supply controlling device in accordance with the present invention can be arranged so that the query generating means can correct the query in accordance with the level of the search result.

According to the arrangement above, provided that the level of the search result determined by the search result comparing means is, for instance, an amateur level, the query generating means modifies the query (exclusive query, for instance) to make the search result have, for instance, a standard level. Also, provided that the level of the search result determined by the search result comparing means is, for instance, a standard level, the query generating means modifies the query to make the search result have, for instance, an expert level. In this manner, the modification of the query in accordance with the level of the search result by the query generating means enables the user to acquire a search result in a higher level than that before the modification.

The invention being thus described, it will be obvious that the same may be varied in many ways. Such variations are not to be regarded as a departure from the spirit and scope of the invention, and all such modifications as would be obvious to one skilled in the art intended to be included within the scope of the following claims.

What is claimed is:

1. A data supply controlling device, comprising:

user data storage means for storing user data for each user who is to be supplied with information through a user terminal;

query generating means for generating, based on at least the user data, a query for a data supplying device to search for the information;

search result storage means for storing a result of a search conducted by the data supplying device for the information in accordance with the query;

identifying data management means for managing identifying data by which each user can be identified, separately from user specifying data by which each user can be specified; and

data supplying means for generating a search result matched with the identifying data of the user from the search result stored in the search result storage means for supply to the user terminal.

2. The data supply controlling device as set forth in claim 1, further comprising

user specifying data management means for managing the user specifying data in such a manner that no access can be made to the user specifying data over a communications network.

3. The data supply controlling device as set forth in claim 1, wherein

the user specifying data is at least one item selected from the group consisting of a name, an address, and an Internet Protocol address of the user.

4. The data supply controlling device as set forth in claim 1, wherein

the identifying data includes a user ID and a certification number.

5. The data supply controlling device as set forth in claim 1, further comprising

query management means for merging and managing the query for each user generated by the query generating means, so as to supply the merged query to the data supplying device.

6. The data supply controlling device as set forth in claim 5, wherein

the query management means merges the query by combining queries sharing an identical keyword to a single query.

7. The data supply controlling device as set forth in claim 5, wherein

the query management means merges the query by deleting a query included in another query.

8. The data supply controlling device as set forth in claim 5, wherein

the data supplying means extracts a search result matched with the identifying data of the user from a search result derived based on the query merged by the query management means for supply to the user terminal.

9. The data supply controlling device as set forth in claim 1, further comprising

user data management means for encrypting and managing the user data in relation to matched identifying data.

10. The data supply controlling device as set forth in claim 1, wherein

the query generating means generates, as the query, a first query to include information necessary to the user in a search result and a second query to conceal the first query, both for supply to the data supplying device.

11. The data supply controlling device as set forth in claim 10, wherein

the first query is composed of a base query to include information matched with the user data in a search result, an additional query to include information excluded by the base query in an information search in the search result, and an exclusive query to exclude each piece of poorly related information from the search result.

12. The data supply controlling device as set forth in claim 10, wherein

the second query is a dummy query to include information unnecessary to the user in the search result.

13. The data supply controlling device as set forth in claim 12, wherein:

the query generating means generates a filtering query to filter out a search result derived from the dummy query from a search result stored in the search result storage means; and

the data supplying means supplies the search result not to be filtered out to the user terminal.

14. The data supply controlling device as set forth in claim 1, wherein

the query generating means calculates an evaluation value E for each of m queries by dividing a sum of L evaluation values which represent evaluation by the user on L past search results by L.

15. The data supply controlling device as set forth in claim 14, wherein

the query generating means employs one of the queries with a high evaluation value E as a base query to include information matched with the user data in a search result.

16. The data supply controlling device as set forth in claim 15, wherein

the query generating means calculates the evaluation value E and selects the base query periodically for each of the m queries.

17. The data supply controlling device as set forth in claim 1, wherein

the data supplying means supplies a search result suitable to each user to the user terminal in response to a data supply request sent from the user through the user terminal.

18. The data supply controlling device as set forth in claim 1, further comprising

search rule storage means for storing a search rule specifying a rule by which the query generating means generates a query in accordance with kinds of information to be searched for,

wherein

the query generating means generates a query based on the user data and the search rule.

19. The data supply controlling device as set forth in claim 18, wherein

the search rule stored in the search rule storage means is supplied from the data supplying device and is updated upon every supply.

20. The data supply controlling device as set forth in claim 1, further comprising:

evaluation value storage means for storing an evaluation value for each piece of information to be searched for in accordance with the user data; and

user data evaluation means for evaluating the user data based on a sum of all the evaluation values and a sum of the evaluation values matched with the user data.

21. The data supply controlling device as set forth in claim 20, wherein

the user data evaluation means evaluates the user data by determining a knowledge level of the user about information to be searched for.

22. The data supply controlling device as set forth in claim 20, wherein:

the user data indicates a degree of ownership of goods given by the information to be searched for by the user; and

the user data evaluation means evaluates the user data by determining an ownership ratio of the goods by the user.

23. The data supply controlling device as set forth in claim 22, wherein

the user data evaluation means evaluates the user data by determining the ownership ratio by the user by giving different weights to different kinds of goods.

24. The data supply controlling device as set forth in claim 22, wherein

the user data evaluation means evaluates the user data by giving a high mark to the user data if TCa=Pa/TPa is high,

where TPa is a sum of evaluation values ax(a, k) (k=1, 2, . . . , n) given to n items of goods belonging to a category a, and Pa is a sum of evaluation values px(a, L) (L=1, 2, . . . , m) given to m items of the goods belonging to the category a and owned by the user.

25. The data supply controlling device as set forth in claim 22, wherein

the goods are books.

26. The data supply controlling device as set forth in claim 1, further comprising

search result comparing means for comparing the search result stored in the search result storage means with a search result suitable to the user, so as to determine a level of the search result suitable to the user.

27. The data supply controlling device as set forth in claim 26, further comprising

query management means for merging and managing the query generated for each user by the query generating means, for supply to the data supplying device,

wherein

the search result comparing means uses a result of a search conducted based on the merged query as the search result stored in the search result storage means for use in a comparison with a search result suitable to each user.

28. The data supply controlling device as set forth in claim 26, wherein

the level is selected from multiple levels of expertise.

29. The data supply controlling device as set forth in claim 28, wherein

the multiple levels include an "amateur" level, a "standard" level, and an "expert" level in ascending order to expertise.

30. The data supply controlling device as set forth in claim 26, wherein

the query generating means corrects the query in accordance with the level of the search result.

31. The data supply controlling device as set forth in claim 30, wherein

the level is selected from multiple levels of expertise; and

the query generating means corrects the query so that the selected level is changeable from one to another among the multiple levels.

32. The data supply controlling device as set forth in claim 31, wherein

the query generating means corrects the query so that the selected level is changeable to a higher level of expertise among the multiple levels.

33. A data supply system, comprising:

a data supply controlling device including: user data storage means for storing user data for each user who is to be supplied with information through a user terminal; query generating means for generating, based on at least the user data, a query for a data supplying device to search for the information; search result storage means for storing a result of a search conducted by the data supplying device for the information in accordance with the query; identifying data management means for managing identifying data by which each user can be identified, separately from user specifying data by which each user can be specified; and data supplying means for generating a search result matched with the identifying data of the user from the search result stored in the search result storage means for supply to the user terminal;

a data supplying device for conducting a search for the information in accordance with a query given by the data supply controlling device to provide a search result to the data supply controlling device; and

a user terminal for receiving a search result suitable to each user from the data supply controlling device,

the data supply controlling device, the data supplying device, and the user terminal being interconnected through a communications network.

34. A data supply controlling device, comprising:

user data storage means for storing user data for each user who is to be supplied with information through a user terminal;

query generating means for generating, based on at least the user data, a query to search data supplied from a data supplying device;

search means for searching for the information in accordance with the query;

search result storage means for storing a result of a search conducted by the search means;

identifying data management means for managing identifying data by which each user can be identified, separately from user specifying data by which each user can be specified; and

data supplying means for generating a search result matched with the identifying data of the user from the search result stored in the search result storage means for supply to the user terminal.

35. A data supply system, comprising:

a data supply controlling device including: user data storage means for storing user data for each user who is to be supplied with information through a user terminal; query generating means for generating, based on at least the user data, a query to search data supplied from a data supplying device; search means for searching for the information in accordance with the query; search result storage means for storing a result of a search conducted by the search means; identifying data management means for managing identifying data by which each user can be identified, separately from user specifying data by which each user can be specified; and data supplying means for generating a search result matched with the identifying data of the user from the search result stored in the search result storage means for supply to the user terminal;

a data supplying device for supplying, to the data supply controlling device, data to be searched by the data supply controlling device; and

a user terminal for receiving a search result suitable to each user from the data supply controlling device,

the data supply controlling device, the data supplying device, and the user terminal being interconnected through a communications network.

36. A method of supplying data, comprising the steps of:

storing, in user data storage means, user data for each user who is to be supplied with information through a user terminal;

generating, based on at least the user data, a query for a data supplying device to search for the information;

storing, in search result storage means, a result of a search conducted by the data supplying device for the information in accordance with the query;

registering, in identifying data management means, identifying data by which each user can be identified, separately from user specifying data by which each user can be specified; and

generating a search result matched with the identifying data of the user from the search result stored in the search result storage means for supply to the user terminal.

37. A method of supplying data, comprising the steps of:

storing, in user data storage means, user data for each user who is to be supplied with information through a user terminal;

generating, based on at least the user data, a query to search data supplied from a data supplying device;

searching for information in accordance with the query;

storing a search result in search result storage means;

registering, in identifying data management means, identifying data by which each user can be identified, separately from user specifying data by which each user can be specified; and

generating a search result matched with the identifying data of the user from the search result stored in the search result storage means for supply to the user terminal.

38. A storage medium for storing a data supply program executed by a computer to implement a method of supplying data, the method comprising the steps of:

storing, in user data storage means, user data for each user who is to be supplied with information through a user terminal;

generating, based on at least the user data, a query for a data supplying device to search for the information;

storing, in search result storage means, a result of a search conducted by the data supplying device for the information in accordance with the query;

registering, in identifying data management means, identifying data by which each user can be identified, separately from user specifying data by which each user can be specified; and

generating a search result matched with the identifying data of the user from the search result stored in the search result storage means for supply to the user terminal.

39. A storage medium for storing a data supply program executed by a computer to implement a method of supplying data, the method comprising the steps of:

storing, in user data storage means, user data for each user who is to be supplied with information through a user terminal;

generating, based on at least the user data, a query to search data supplied from a data supplying device;

searching for information in accordance with the query;

storing a search result in search result storage means;

registering, in identifying data management means, identifying data by which each user can be identified, separately from user specifying data by which each user can be specified; and

generating a search result matched with the identifying data of the user from the search result stored in the search result storage means for supply to the user terminal.

* * * * *

US006606618B2

(12) **United States Patent**
Delo

(10) **Patent No.: US 6,606,618 B2**
(45) **Date of Patent: Aug. 12, 2003**

(54) **METHOD FOR OPTIMIZING THE PERFORMANCE OF A DATABASE**

(75) Inventor: **John C. Delo**, Bellevue, WA (US)

(73) Assignee: **Microsoft Corporation**, Redmond, WA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 106 days.

(21) Appl. No.: **09/792,661**

(22) Filed: **Feb. 23, 2001**

(65) **Prior Publication Data**

US 2001/0032199 A1 Oct. 18, 2001

Related U.S. Application Data

(62) Division of application No. 09/157,853, filed on Sep. 21, 1998.

(51) **Int. Cl.**[7] **G06F 17/30**

(52) **U.S. Cl.** **707/2**; 707/3; 707/10; 707/100; 707/201

(58) **Field of Search** ..... 707/1, 2, 3, 4, 707/5, 100, 101, 102, 103 R, 104.1, 201, 203, 10

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,410,704	A *	4/1995	Norden-Paul et al.	395/700
5,675,784	A *	10/1997	Maxwell et al.	395/611
5,717,924	A *	2/1998	Kawai	395/613
5,752,018	A *	5/1998	Sheffield	395/602
5,864,842	A *	1/1999	Pederson et al.	707/3
5,918,225	A *	6/1999	White et al.	707/3
5,930,779	A *	7/1999	Knoblock et al.	705/412
6,032,151	A *	2/2000	Arnold et al.	707/103
6,044,369	A *	3/2000	Black	707/4
6,058,389	A *	5/2000	Chandra et al.	707/1

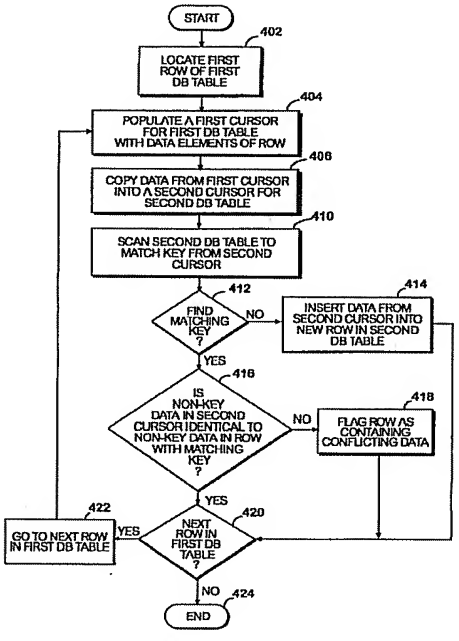
* cited by examiner

Primary Examiner—Hosain T. Alam
Assistant Examiner—Jacques Veillard
(74) *Attorney, Agent, or Firm*—Workman, Nydegger & Seeley

(57) **ABSTRACT**

A relational installation database for storing data elements in the form of strings, objects, etc. is aliased with integer identifiers corresponding to each data element is disclosed. The integer identifiers are obtained from an index that sequentially stores a copy of each unique occurrences of a data element. Populating an installation database with only integers reduces persistent size and provides uniformity to the data fields underlying the database tables, and provides a significant improvement in database performance The uniform data fields may be expanded and contracted to add temporary rows and columns directly to a database table. Database tables may be created in a modular fashion and may be efficiently merged together when the software product is complete. Also, differences between various versions of the software product may be recorded in database transforms.

18 Claims, 6 Drawing Sheets



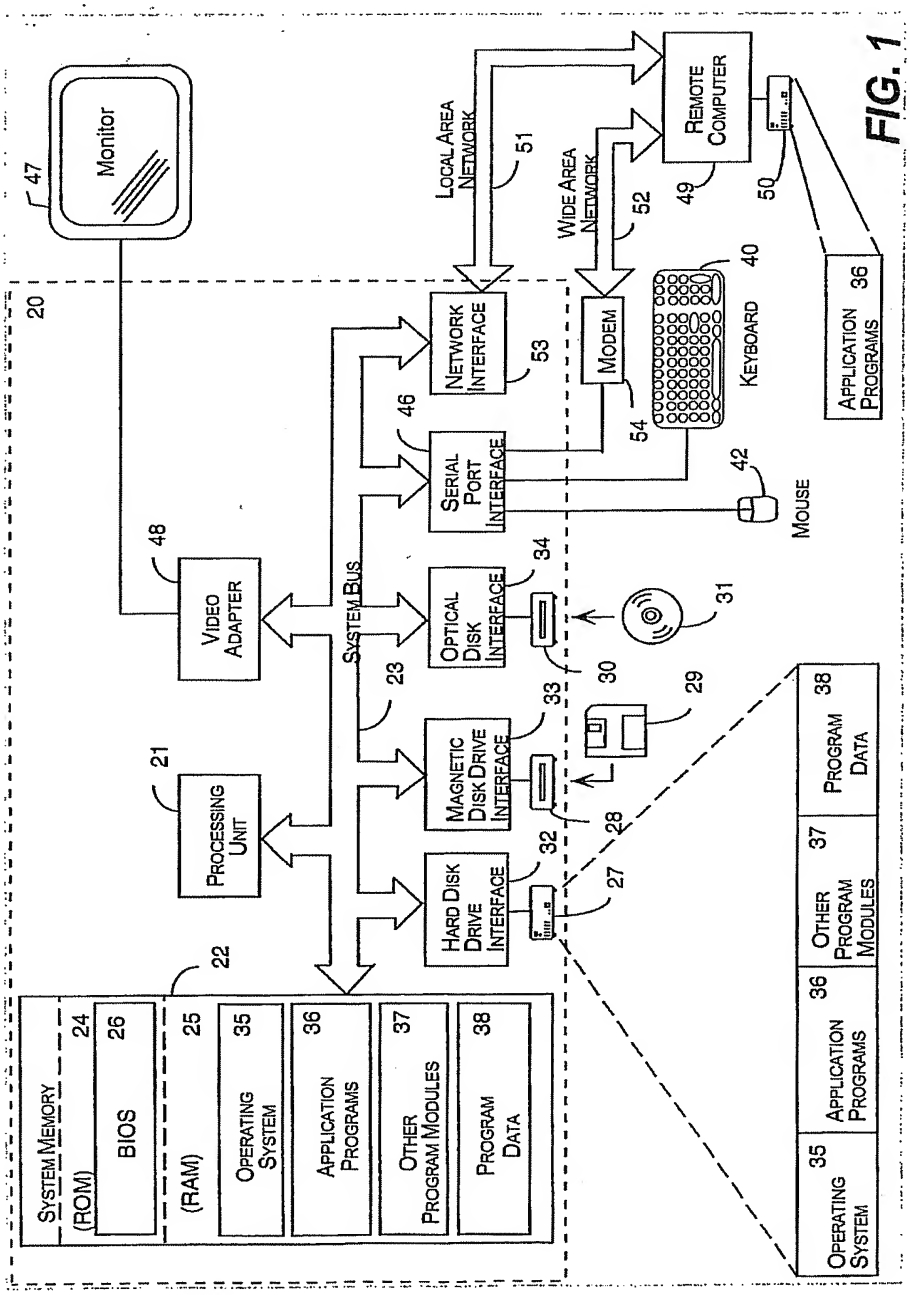

Monitor
47
20
Local Area Network
51
Wide Area Network
52
Remote Computer
49
50
Application Programs
36
Keyboard
40
Modem
54
Network Interface
53
Serial Port Interface
46
Mouse
42
Video Adapter
48
System Bus
23
Optical Disk Interface
34
31
30
Magnetic Disk Drive Interface
33
29
28
Hard Disk Drive Interface
32
27
Processing Unit
21
22
System Memory
24
(ROM)
BIOS
26
(RAM)
25
Operating System
35
Application Programs
36
Other Program Modules
37
Program Data
38

FIG. 1

200

207 205 209 211 212 213

Feature_Table

Feature_ID	Feature_Name	Attribute_F_A	Attribute_F_B	Feature_Preference
Feature_1	Feature_One	Attribute_F_One	Attribute_F_Two	1
Feature_2	Feature_Two	Attribute_F_Three	Attribute_F_One	1
Feature_3	Feature_Three	Attribute_F_Five	Attribute_F_Four	1

217 215 219 221 222 223

Component_Table

Component_ID	Component_Name	Attribute_C_A	Attribute_C_B	Component_Preference
Component_1	Component_One	Attribute_C_Three	Attribute_One	1
Component_2	Component_Two	Attribute_C_One	Attribute_Two	0
Component_3	Component_Three	Attribute_C_Four	Attribute_Five	1
Component_4	Component_Four	Attribute_C_Three	Attribute_Six	2
Component_5	Component_Five	Attribute_C_One	Attribute_Seven	1
Component_6	Component_Six	Attribute_C_Eight	Attribute_Two	2
Component_7	Component_Seven	Attribute_C_Nine	Attribute_Nine	3
Component_8	Component_Eight	Attribute_C_Three	Attribute_ten	3

225

Feature_Comp_Table

207 Feature_ID	217 Component_ID
Feature_1	Component_1
Feature_1	Component_2
Feature_1	Component_3
Feature_2	Component_4
Feature_2	Component_5
Feature_2	Component_6
Feature_3	Component_7
Feature_3	Component_7

PRIOR ART

FIG.2A

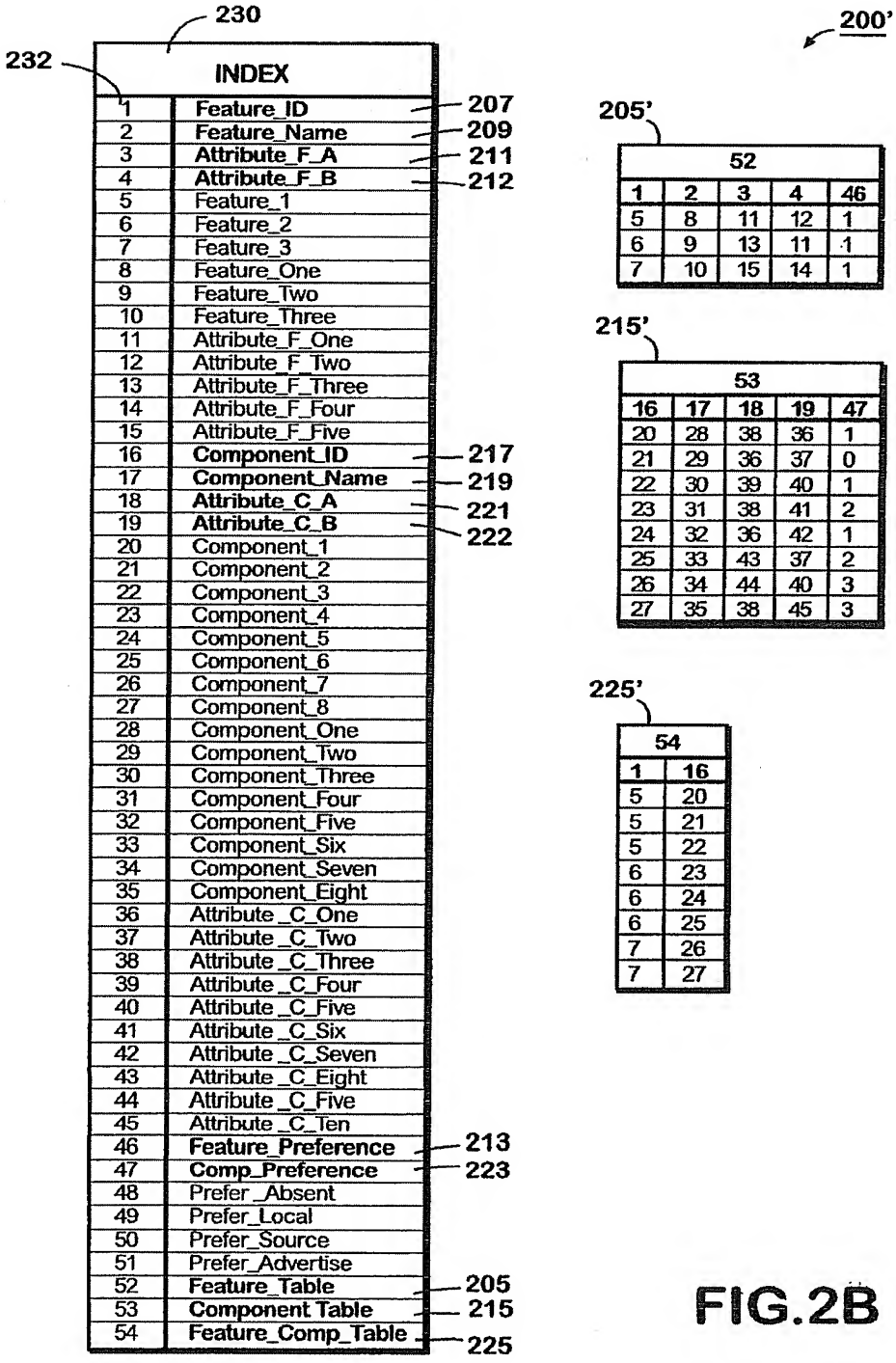

230
232
INDEX
1 Feature_ID
2 Feature_Name
3 Attribute_F_A
4 Attribute_F_B
5 Feature_1
6 Feature_2
7 Feature_3
8 Feature_One
9 Feature_Two
10 Feature_Three
11 Attribute_F_One
12 Attribute_F_Two
13 Attribute_F_Three
14 Attribute_F_Four
15 Attribute_F_Five
16 Component_ID
17 Component_Name
18 Attribute_C_A
19 Attribute_C_B
20 Component_1
21 Component_2
22 Component_3
23 Component_4
24 Component_5
25 Component_6
26 Component_7
27 Component_8
28 Component_One
29 Component_Two
30 Component_Three
31 Component_Four
32 Component_Five
33 Component_Six
34 Component_Seven
35 Component_Eight
36 Attribute_C_One
37 Attribute_C_Two
38 Attribute_C_Three
39 Attribute_C_Four
40 Attribute_C_Five
41 Attribute_C_Six
42 Attribute_C_Seven
43 Attribute_C_Eight
44 Attribute_C_Five
45 Attribute_C_Ten
46 Feature_Preference
47 Comp_Preference
48 Prefer_Absent
49 Prefer_Local
50 Prefer_Source
51 Prefer_Advertise
52 Feature_Table
53 Component Table
54 Feature_Comp_Table
207
209
211
212
217
219
221
222
213
223
205
215
225
200'
205'
52
1 2 3 4 46
5 8 11 12 1
6 9 13 11 1
7 10 15 14 1
215'
53
16 17 18 19 47
20 28 38 36 1
21 29 36 37 0
22 30 39 40 1
23 31 38 41 2
24 32 36 42 1
25 33 43 37 2
26 34 44 40 3
27 35 38 45 3
225'
54
1 16
5 20
5 21
5 22
6 23
6 24
6 25
7 26
7 27

FIG.2B

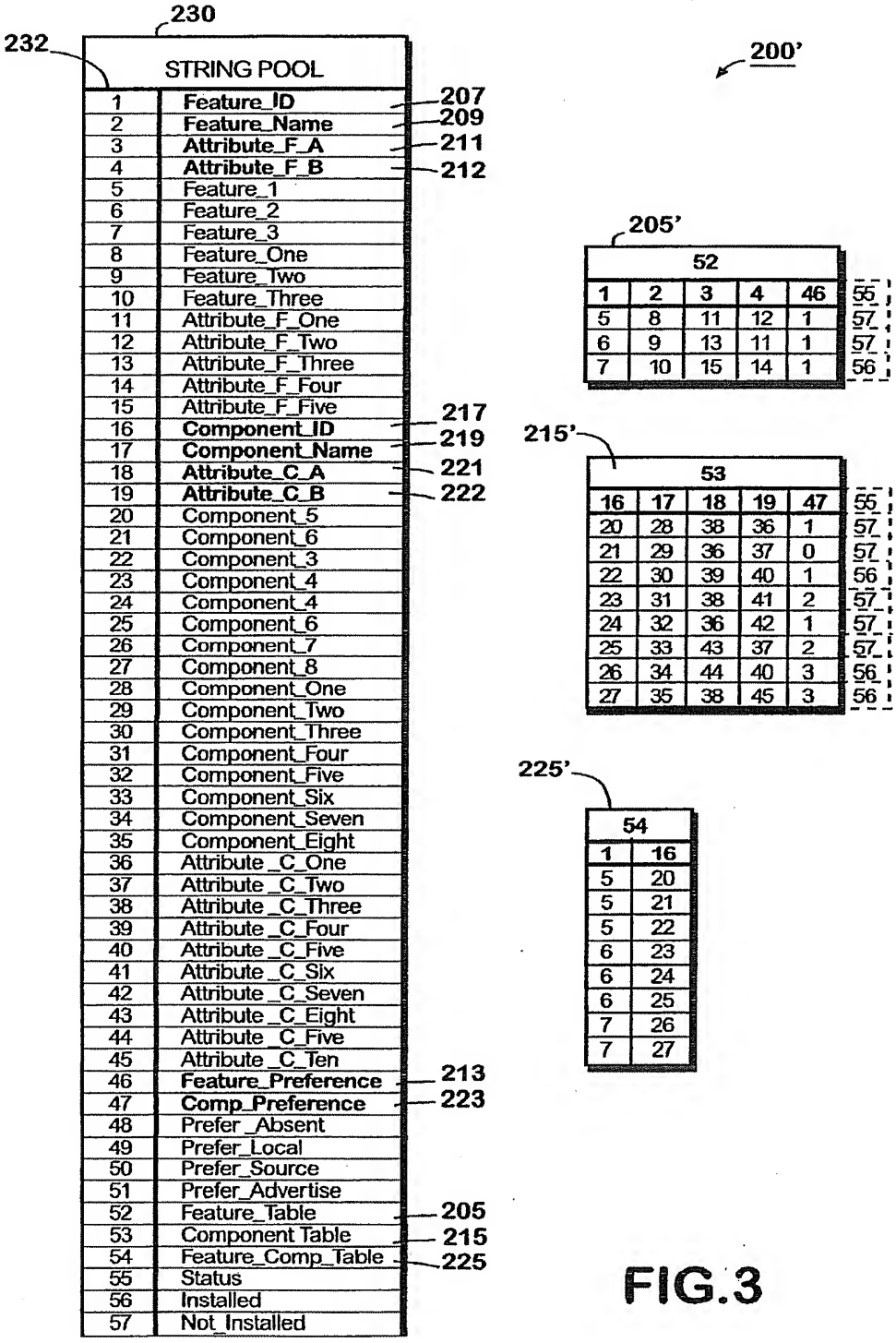

230
232
STRING POOL
1 Feature_ID
2 Feature_Name
3 Attribute_F_A
4 Attribute_F_B
5 Feature_1
6 Feature_2
7 Feature_3
8 Feature_One
9 Feature_Two
10 Feature_Three
11 Attribute_F_One
12 Attribute_F_Two
13 Attribute_F_Three
14 Attribute_F_Four
15 Attribute_F_Five
16 Component_ID
17 Component_Name
18 Attribute_C_A
19 Attribute_C_B
20 Component_5
21 Component_6
22 Component_3
23 Component_4
24 Component_4
25 Component_6
26 Component_7
27 Component_8
28 Component_One
29 Component_Two
30 Component_Three
31 Component_Four
32 Component_Five
33 Component_Six
34 Component_Seven
35 Component_Eight
36 Attribute _C_One
37 Attribute _C_Two
38 Attribute _C_Three
39 Attribute _C_Four
40 Attribute _C_Five
41 Attribute _C_Six
42 Attribute _C_Seven
43 Attribute _C_Eight
44 Attribute _C_Five
45 Attribute _C_Ten
46 Feature_Preference
47 Comp_Preference
48 Prefer _Absent
49 Prefer_Local
50 Prefer_Source
51 Prefer_Advertise
52 Feature_Table
53 Component Table
54 Feature_Comp_Table
55 Status
56 Installed
57 Not_Installed
207
209
211
212
217
219
221
222
213
223
205
215
225
200'
205'
52
1 2 3 4 46 55
5 8 11 12 1 57
6 9 13 11 1 57
7 10 15 14 1 56
215'
53
16 17 18 19 47 55
20 28 38 36 1 57
21 29 36 37 0 57
22 30 39 40 1 56
23 31 38 41 2 57
24 32 36 42 1 57
25 33 43 37 2 57
26 34 44 40 3 56
27 35 38 45 3 56
225'
54
1 16
5 20
5 21
5 22
6 23
6 24
6 25
7 26
7 27

FIG.3

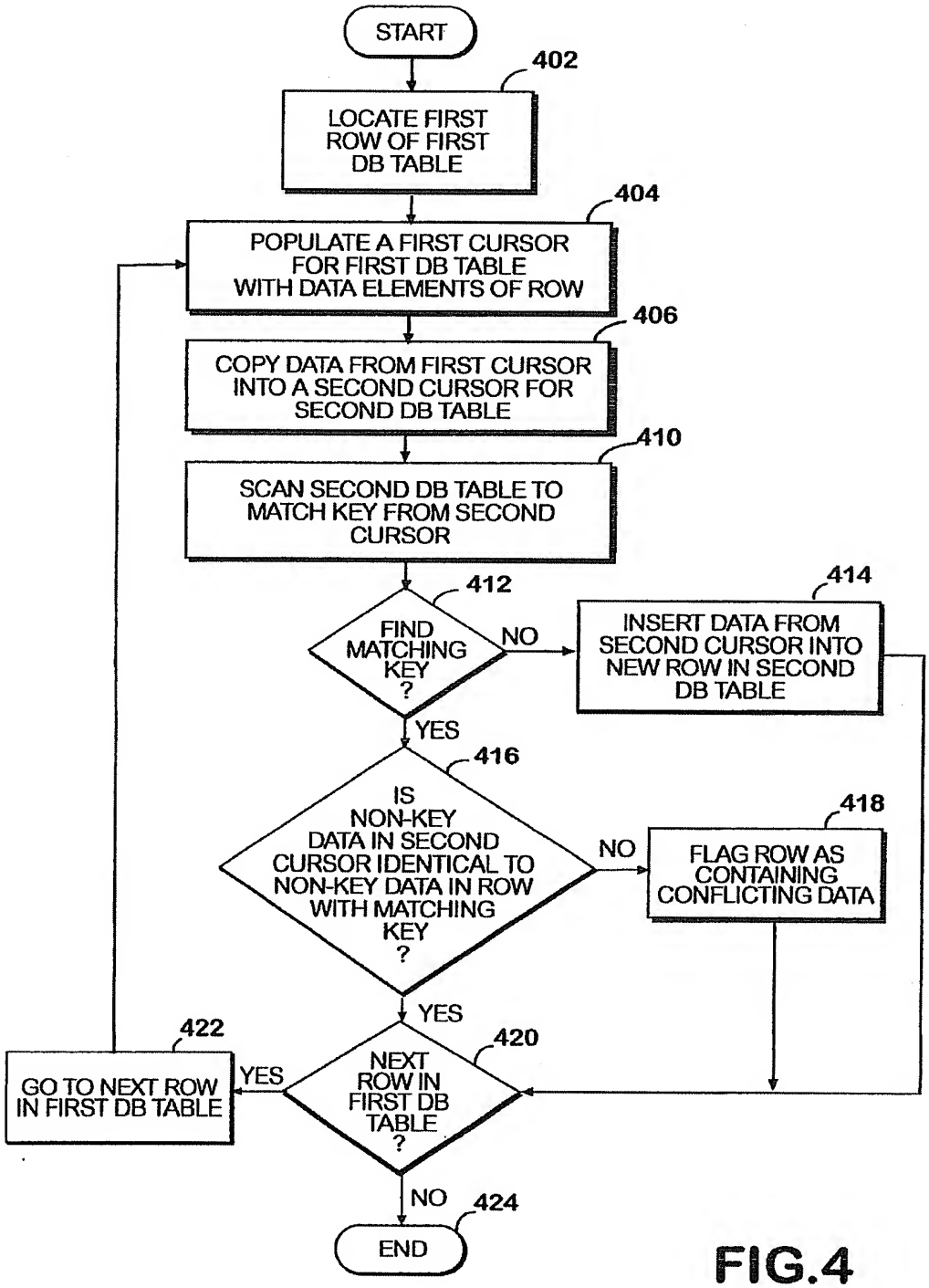

START
402
LOCATE FIRST ROW OF FIRST DB TABLE
404
POPULATE A FIRST CURSOR FOR FIRST DB TABLE WITH DATA ELEMENTS OF ROW
406
COPY DATA FROM FIRST CURSOR INTO A SECOND CURSOR FOR SECOND DB TABLE
410
SCAN SECOND DB TABLE TO MATCH KEY FROM SECOND CURSOR
412
FIND MATCHING KEY ?
NO
414
INSERT DATA FROM SECOND CURSOR INTO NEW ROW IN SECOND DB TABLE
YES
416
IS NON-KEY DATA IN SECOND CURSOR IDENTICAL TO NON-KEY DATA IN ROW WITH MATCHING KEY ?
NO
418
FLAG ROW AS CONTAINING CONFLICTING DATA
YES
420
NEXT ROW IN FIRST DB TABLE ?
YES
422
GO TO NEXT ROW IN FIRST DB TABLE
NO
424
END

FIG.4

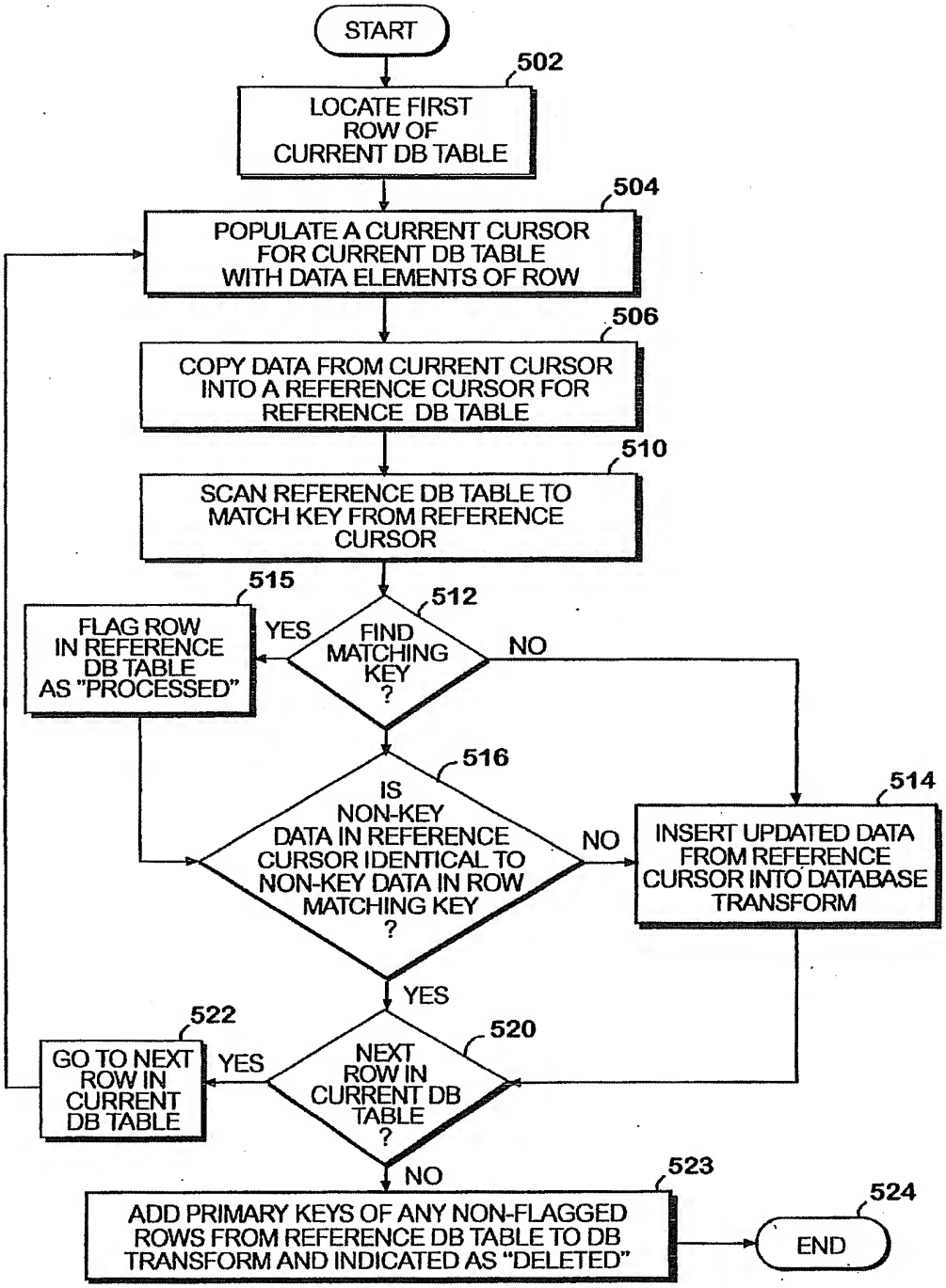

START
502
LOCATE FIRST ROW OF CURRENT DB TABLE
504
POPULATE A CURRENT CURSOR FOR CURRENT DB TABLE WITH DATA ELEMENTS OF ROW
506
COPY DATA FROM CURRENT CURSOR INTO A REFERENCE CURSOR FOR REFERENCE DB TABLE
510
SCAN REFERENCE DB TABLE TO MATCH KEY FROM REFERENCE CURSOR
512
FIND MATCHING KEY ?
YES
515
FLAG ROW IN REFERENCE DB TABLE AS "PROCESSED"
NO
516
IS NON-KEY DATA IN REFERENCE CURSOR IDENTICAL TO NON-KEY DATA IN ROW MATCHING KEY ?
NO
514
INSERT UPDATED DATA FROM REFERENCE CURSOR INTO DATABASE TRANSFORM
YES
520
NEXT ROW IN CURRENT DB TABLE ?
YES
522
GO TO NEXT ROW IN CURRENT DB TABLE
NO
523
ADD PRIMARY KEYS OF ANY NON-FLAGGED ROWS FROM REFERENCE DB TABLE TO DB TRANSFORM AND INDICATED AS "DELETED"
524
END

FIG.5

METHOD FOR OPTIMIZING THE PERFORMANCE OF A DATABASE

RELATED APPLICATIONS

This application is a divisional of application No. Ser. 09/157,853 filed Sep. 21, 1998, which is related to the following applications, all of which were filed on the same day and assigned to the same assignee as the parent application:

"Use of Relational Databases for Software Installation," now U.S. Pat. No. 6,237,144;

"System and Method for Repairing a Damaged Application Program," now U.S. Pat. No. 6,397,381;

"Method and System for Restoring a Computer to its Original State After an Unsuccessful Installation Attempt," now U.S. Pat. No. 6,363,499;

"A Method for Categorizing and Installing Selected Software Components," now U.S. Pat. No, 6,370,686;

"System and Method for Managing Locations of Software Components Via a Source List," now U.S. Pat. No, 6,269,377;

"Software Installation and Validation Using Custom Actions," now U.S. Pat. No. 6,378,127;

"Internal Database Validation," now U.S. Pat. No. 6,389,414;

"Management of Non-persistent Data in a Persistent Database"—Ser. No. 09/157,883;

"Method and System for Advertising Applications," now U.S. Pat. No. 6,345,386; and

"Software Implementation Installer Mechanism," now U.S. Pat. No. 6,418,554.

TECHNICAL FIELD

The present invention generally relates to performance optimizations for relational databases. for the process of installing a software product onto a target computer system. More particularly, the present invention relates to performance optimizations for an installation database system for installing a software product onto a target computer system.

BACKGROUND OF THE INVENTION

A software product may be provided to an end-user through a variety of media. For example, a software product may be supplied via a magnetic disk, an optical disk or a network connection. The end-user may install the software product onto a target computer system by copying various components of the software product into the memory of the target computer system. To facilitate the installation process, the media containing the software product may include an installation database system that describes all of the features and components of the software product, as well as any additional installation information that may be helpful to the target computer system.

A relational database model may serve as an effective installation database system due to its ability to provide fast and efficient query processing, minimal redundancy of data and adequate data integrity. A relational database comprises a collection of data elements stored in a plurality of two-dimensional database tables. A row of a relational database table defines a data record, whereas each column in a row defines a single data element. Data elements may comprise integer data, string data and even simple object data.

The relational database model maintains a rigid storage structure, wherein the data fields comprising the columns of a relational database table are of a predetermined fixed length. As such, each row in a relational database table is of a uniform size. This uniform storage model allows for simple memory management techniques and minimal CPU cycles when a data record (row) is fetched from disk. Once a data record is fetched into main memory, a database engine can readily read or write a particular data element by simply jumping to a known offset for a given column in the data record and then reading or writing the next n bytes, where n is the size of a column in number of characters.

On the downside, however, a rigid storage structure may cause the persistent size of a relational database to be very large. For example, data elements that are too large to fit within in a fixed-length data field must be truncated. Thus, in order to avoid loss of data, data fields are typically designed to be large enough to accommodate the longest data element that is likely to be stored in the column. If a given data element does not occupy all the space available in the data field, the unoccupied space is filled with null values. Therefore, the rigid storage structure of the relational database model may lead to a great deal of wasted memory space and cause the performance of the database system to be diminished.

Also, the rigid storage structure of the relational database model does not allow for the direct addition of temporary rows and columns to a database table. Temporary rows and columns may be useful in managing data, such as installation data, that is dependant on dynamic external conditions.

An installation database is generally created and populated by the software developer that supplies the software product. For each version of the software product, the software developer must create and populate a new installation database. Presently, if an end-user desires to add enhancements, patches, or upgrades to a software product, a complete version of a new installation database must be obtained to replace the original installation database.

Accordingly, there is a need for a relational installation database model that maintains a uniform data storage structure while eliminating wasted memory space. There is also a need for a relational database model that provides for the direct addition of temporary rows and columns to a database tables. There is a further need for a method of modifying an installation database or an installed prior version of a software product to reflect upgrades, patches, or enhancements without supplying a complete version of a new installation database.

SUMMARY OF THE INVENTION

The present invention fulfills the needs in the art by providing optimizations for the process involved in the installation of a software product onto a target computer system. The present invention provides a relational installation database system that has increased performance and reduced persistent size. Persistent size is reduced by substituting non-integer data elements in a database table, whether in the form of strings, objects, etc., with integer identifiers. The integer identifiers are obtained from an index which maintains a sequential record of each unique occurrences of a data element. In effect, the data elements in a database table are aliased with integer identifiers and the relationship between the integer identifiers and the data elements is determined by accessing the index.

The data fields underlying a database table are two-dimensional data arrays. Aliasing the non-integer data elements of a database table with integers causes the data arrays to be uniform. Thus, manipulation of the data arrays is simplified. The present invention provides for the direct

addition of temporary rows and columns to a database table by expanding and contracting the underlying data arrays. When temporary rows or columns are created, data elements in both the temporary and permanent rows and columns may be accessed and modified. Data elements in the permanent rows and columns persist in the database, while data elements in the temporary rows and columns do not persist.

Functionality is provided to allow installation database tables to be merged together. This allows software developers to develop software in a modular fashion and to create an installation database for each module. Installation database may then be merged into a final installation database that is shipped to the end-user. Further, differences between two or more installation database tables may be recorded in a database transform. This database transform may be applied to a database table in such a way that the changes are incorporated therein. Thus, software developers are able to provide enhancements, upgrades, patches, etc. to an end-user in the form of database transform, without the need for shipping an entire modified installation database.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** is a functional block diagram of a computer system that provides an operating environment for the exemplary embodiments of the present invention;

FIG. **2A** illustrates a traditional relational database;

FIG. **2B** illustrates a relational database that has been aliased with integer identifiers;

FIG. **3** illustrates the direct addition of a temporary column to a relational database table;

FIG. **4** is a flow chart demonstrating the process involved in an exemplary merger of two installation database tables; and

FIG. **5** is a flow chart demonstrating the process involved in creating a transform database table to record the differences between two installation database tables.

DETAILED DESCRIPTION OF THE EXEMPLARY EMBODIMENTS

The present invention provides optimizations for the process of installing a software product onto a target computer system. In particular, the present invention contemplates the use of an installation database system to facilitate the installation process. As used herein, the phrase "database system" is meant to indicate a database and any associated database engines or management systems. An installation database may be a relational database for storing data elements that describe the features and components of a software product as well as other installation information. The present invention serves to optimize the performance of an installation database system by reducing persistent size and allowing greater flexibility in the physical structure of the installation database. The present invention also provides methods for utilizing an installation database to efficiently provide upgrades, patches, enhancements, etc. to an end user.

Still, while the exemplary embodiment of the present invention will be described in the general context of an installation database system, those skilled in the art will appreciate that certain principles of the present invention may be applied to any relational database system, irrespective of the particular application thereof.

Exemplary Operating Environment

Referring now to the drawings, in which like numerals indicate like elements throughout the several figures, FIG. **1** and the following discussion are intended to provide a brief, general description of a suitable computing environment in which the invention may be implemented. While the invention will be described in the general context of an installation database system that is run by an operating system in conjunction with a personal computer, those skilled in the art will recognize that the invention also may be implemented in combination with other program modules. Generally, program modules include routines, programs, data structures, etc. that perform particular tasks or implement particular abstract data types. Moreover, those skilled in the art will appreciate that the invention may be practiced with other computer system configurations, including hand-held devices, multiprocessor systems, microprocessor-based or programmable consumer electronics, minicomputers, mainframe computers, and the like. The invention may also be practiced in distributed computing environments where tasks are performed by remote processing devices that are linked through a communications network. In a distributed computing environment, program modules may be located in both local and remote memory storage devices.

An exemplary operating environment **100** for implementing the invention includes a conventional personal computer system **120**, including a processing unit **121**, a system memory **122**, and a system bus **123** that couples the system memory **122** to the processing unit **121**. The system memory **122** includes read only memory (ROM) **124** and random access memory (RAM) **125**. A basic input/output system **126** (BIOS), containing the basic routines that help to transfer information between elements within the personal computer system **120**, such as during start-up, is stored in ROM **124**.

The personal computer system **120** further includes a hard disk drive **127**, a magnetic disk drive **128**, e.g., to read from or write to a removable magnetic disk **129**, and an optical disk drive **130**, e.g., for reading a CD-ROM disk **131** or to read from or write to other optical media. The hard disk drive **127**, magnetic disk drive **128**, and optical disk drive **130** are connected to the system bus **123** by a hard disk drive interface **132**, a magnetic disk drive interface **133**, and an optical drive interface **134**, respectively. The drives and their associated computer-readable media provide nonvolatile storage for the personal computer system **120**. Although the description of computer-readable media above refers to a hard disk, a removable magnetic disk and a CD-ROM disk, it should be appreciated by those skilled in the art that other types of media that are readable by a computer system, such as magnetic cassettes, flash memory cards, digital video disks, Bernoulli cartridges, and the like, may also be used in the exemplary operating environment.

The computer system **120** may include additional input devices (not shown), such as a microphone, joystick, game pad, satellite dish, scanner, or the like. These and other input devices are often connected to the processing unit **121** through a serial port interface **146** that is coupled to the system bus, but may be connected by other interfaces, such as a game port or a universal serial bus (USB). A monitor **147** or other type of display device is also connected to the system bus **123** via an interface, such as a video adapter **148**. In addition to the monitor, personal computer systems typically include other peripheral output devices (not shown), such as speakers or printers.

The personal computer system **120** may operate in a networked environment using logical connections to one or more remote computer systems, such as a remote computer system **149**. The remote computer system **149** may be a server, a router, a peer device or other common network

node, and typically includes many or all of the elements described relative to the personal computer system **120**, although only a memory storage device **150** has been illustrated in FIG. **1**. The logical connections depicted in FIG. **1** include a local area network (LAN) **151** and a wide area network (WAN) **152**. Such networking environments are commonplace in offices, enterprise-wide computer networks, intranets and the Internet.

When used in a LAN networking environment, the personal computer system **120** is connected to the LAN **151** through a network interface **153**. When used in a WAN networking environment, the personal computer system **120** typically includes a modem **154** or other means for establishing communications over the WAN **152**, such as the Internet. The modem **154**, which may be internal or external, is connected to the system bus **123** via the serial port interface **146**. In a networked environment, program modules depicted relative to the personal computer system **120**, or portions thereof, may be stored in the remote memory storage device. It will be appreciated that the network connections shown are exemplary and other means of establishing a communications link between the computer systems may be used. It will be further appreciated that the invention could equivalently be implemented on host or server computer systems other than personal computer systems, and could equivalently be transmitted to the host computer system by means other than a CD-ROM, for example, by way of the network connection interface **153**.

A number of program modules may be stored in the drives and RAM **125** of the computer system **120**, including an operating system **135**, an installation program module **137** and other program modules **138**. In particular, the installation program module **137** may comprise a database engine or database management system for interacting with an installation database **139**. The purpose of the installation program module **137** is to install the software product **136** onto a target computer system **120**. The software product **136** and installation database **139** are generally provided via a computer-readable medium, such as a CD-ROM disk **131**.

An installation database **139** is a file containing all the installation information (data) for a particular software product **136** that is to be installed onto a target computer system **120**. The binary images comprising the software product **136** may be stored in a location external to the installation database **139**, e.g. in distinct directories on the CD-ROM disk **131**. Alternatively, the binary images may be stored in compressed cabinet files contained in streams within the installation database **120**.

A CD-ROM **131** is typically introduced to the target computer system **120** through an optical drive **130**. The software product **136** and installation database may also be provided via a magnetic disk **129**, a memory storage device **150** of a remote computer system **149**, or any other commonly known medium. Accordingly, the software product **136** and the installation database **139** may be introduced via a magnetic disk drive **128**, or transferred from a remote computer **149** through a serial port interface **146** or a network interface **153**. Other media and/or devices for providing and introducing a software product **136** to a target computer system **120** will be apparent to those skilled in the art and are considered to be within the spirit and scope of the present invention.

As shown, the installation program module **137** may reside on the hard disk drive **127** of the target computer system and may operate in conjunction with the operating system **135**. For example, included with the WINDOWS operating system produced by Microsoft Corporation of Redmond, Wash. is a WINDOWS INSTALLER program module that is operable to detect and install a new software product **136** that is introduced to the target computer system **120**. In addition, the installation program module **137** may be provided along with the software product **136** and the installation database **139** via a computer-readable medium, such as CD-ROM **131**, or the like. Those skilled in the art will appreciate that the installation program module **137**, the software product **136** and the installation database **139** may be provided independently of each other.

Once the software product **136**, the installation database **139** and/or the installation program module **137** are introduced to the target computer system **120**, the processing unit **121** executes the instructions of the installation program module **137** to install an install-set of the software product **136**, i.e., the set of components that are to be installed, onto the target computer system **120**. The application program module **136** may be thought of as comprising a set of features, which in turn comprise a set of components. Certain components may be required for proper operation of a feature of the software product **136**, while other components may be optional.

The installation program module **137** may gather information regarding the current state of the computer system **120**. The information gathered by the installation program module **137**, in conjunction with information stored in the installation database **139** is used to determine which components of the software product **136** are to be included in the install-set. The install-set typically includes all required components and any optional components that are determined by the installation program module **137** to be appropriate for installation onto the target computer system **120**.

Operation of an Exemplary Embodiment

As mentioned above, the performance advantages provided by the relational database model are desirable in the context of an installation database. As such, an installation database may be a relational database comprising a collection of relational database tables. Each relational database table is identified by a unique table name. FIG. **2A** shows a sample relational database comprising three tables: a Feature_Table **205**, a Component_Table **215** and a Feature_Comp_Table **225**. Relational database tables are two dimensional arrays comprising rows and columns. The rows of a relational database table define data records and the columns of each row represent individual data elements. Relational database rows, or records, are distinguished from each other through the use of keys. Relational database columns are identified by column names, also referred to as attributes. These and other concepts of relational databases are well known in the art.

In an exemplary embodiment, the data elements stored in the installation database tables **205**, **215** and **225** correspond to software products, features of the software products and components of the features of the software products. As used herein, the phrase "software product" is meant to refer to an application program module or a suite of application program modules. A feature is a granular piece of a software product. The installation of a feature is optional. Features correspond roughly to the functional features of the software product itself, such as a "Proofing Tools" feature or a "Word" feature. Each feature is essentially a grouping of components and may also include other features. Features need not be globally unique, and therefore may be identified by any appropriate means, such as with a textual feature identifier.

A "component" is a collection of resources, such as or registry keys, that are all installed or uninstalled as a unit, Components are the building blocks of the software product that are not exposed to the user, A resource, such as a file or a registry key, may be part of only one component. Two components may not share the same resource whether they are part of the same software product **136** or parts of different software products. A discussion of features and components of a software product, and their role in an installation database, is described in U.S. Pat. No. 6,237,144, entitled "Use of Relational Databases for Software Installation" filed on Sep. 21, 1998 and owned by the assignee for the present application, This related application, U.S. Pat. No. 6,237,144, is hereby fully incorporated herein by reference.

As shown in FIG. 2A, Feature_Table **205** stores a list of features, which are identified by Feature_ID **207**. Each record in the Feature_Table **205** also includes a Feature_Name **209**, an Attribute_F_A **211**, an Attribute_F_B **212** and a Feature_Preference **213**. The Component_Table **215** is constructed in the manner similar to the Feature_Table **205**, but stores records for components instead of features. Each record in the Component_Table **205** includes a Component_Name **219**, an Attribute_C_A **221**, an Attribute_C_B **223** and a Component_Preference **223**. The Feature_Comp_Table **225** associates Feature_IDs **207** with Component_IDs **217**. As can be seen, the data elements stored in the database tables of FIG. 2A comprise string data, each string having a variable length, and integer data.

In order to improve the performance of the database engine within the installation program module **137**, the present invention undertakes to simplify and reduce the physical storage model of the installation database **139**. Such simplification and reduction is accomplished by populating an installation database table with only integer identifiers that represent each of the installation data elements. In effect, each non-integer data element is aliased with an integer identifier. In order to keep track of the original data elements, an index is created for storing a single copy of each unique string, object or other non-integer type of installation data that may be included in the installation database **139**. The data element corresponding to a given integer identifier in a database table may be determined by reference to the index.

Integers are generally more uniform and compact than strings or objects. Thus, by aliasing an installation database with integers, the size and complexity of the database may be significantly reduced. For example, the columns of the installation database **139** may be set to a fixed size due to the single type of data stored therein. Also, less memory is utilized because only one copy of each string or object is stored in the index. As another benefit, the index storing the string and object data elements is not a database table and is thus not subject to any size restrictions that would be imposed on a database table. Thus, for example, strings stored in the index may be up to 64 k bytes in size. In addition to strings, reference-counted objects (COM objects) are also used in the exemplary installation database and thus, object references may be stored directly as integer values when they fit in the size of a database integer (32-bits). The use of COM objects insures that only a single copy of the object exists for any number of references to it. For objects that must be referenced by 64-bit pointers, a separate index may be created and such objects may be aliased in a manner similar to strings.

In addition to keeping the index of data elements, a corresponding array is required to hold a reference count for each data element. A reference count is used to keep track of each data element, so that a data element may be released from memory when it is no longer referenced by any table or cursor in the database. Both the data elements and the corresponding reference counts must be persisted when the database is committed. An additional complexity occurs with temporary data elements, which are discussed in greater detail below. Temporary data elements must operationally behave in the same manner as persistent data, but when the database is committed, the reference counts for the temporary data elements must be excluded, as well as the data elements themselves if they have only temporary references.

FIG. 2B illustrates an installation database **200'** that has been aliased with integer identifiers. Aliased installation database **200'** corresponds to the installation database **200** of FIG. 2A. In FIG. 2B, index **230** stores each unique occurrence of a string data element that is found in the Feature_Table **205**, the Component_Table **215** and the Feature_Comp_Table **225** of installation database **200**. For example, the string "Feature_ID" **207** is stored as a first entry in the index **230**. As can be seen, there are no duplicate entries in the index **230**. The left-most column **232** in the index **230** stores sequential integer values identifying each entry in the index **230**. Column **232** is shown for illustrative purposes only and is actually superfluous because a computer system will automatically assign sequential integers identifiers to the elements of a sequential index structure, such as index **230**.

The sequential integer identifiers created in the index **230** are used to alias the data elements in the database tables. For example, Feature_Table **205** from FIG. 2A may be converted into table "52" **205'** shown in FIG. 2B. As can be seen, entry number **52** in the index **230** corresponds to the string "Feature_Table" **205**. Also, entry number **1** in the index **230** corresponds to the string "Feature_ID." Therefore, the integer identifier "1" is stored as the first entry in table "52" **205'**. Similarly, entry number **5** in the index **230** corresponds to the string "Feature_1." Accordingly, the integer identifier "5" is stored in table 52 **205'** in the location corresponding to the field where the string "Feature_1" is stored in Feature_Table **205**. All rows and columns of each database table in the exemplary installation database are populated in the same fashion. It should be noted that the integer identifiers are stored in the database table columns having a data type other than integer. Thus, in aliased table "52," there is no conflict between the integer identifier "1" and the original integer data stored in column "46". Comparing the database tables **205**, **215** and **225** of FIG. 2A to the corresponding database tables **205'**, **215'** and **225'** of FIG. 2B, it can be seen that the persistent size of database tables **205'**, **215'** and **225'** has been significantly reduced.

In the exemplary embodiment, index **230** is a hashed index. As such, a string data element and its corresponding integer identifier may be located in index **230** in a quick and efficient manner. The installation program module **137** will query the installation database **200'** in order to access or modify string data elements. A database query may include the name of one or more string data elements. Integer identifiers corresponding to the queried string data elements in may be located in the index **230** using any well-known hashing function. The appropriate integer identifiers are then copied into database cursors, which are used to locate the corresponding aliased data elements in the database tables. Database cursors are well known within the art and further discussion thereof is not necessary.

An example database query that might be used in conjunction with the traditional relational database **200** shown in FIG. 2A is as follows:

Select Component_Name

From Component_Table

Where Component_Preference="0".

This standard database query seeks the Component $_{13}$ Name **219** of any component in the Component_Table **215** that has a Component_Preference **223** of "0". The result set for this query comprises the data element "Component_Two." In the aliased database system of FIG. **2B**, the above example query would be effectively converted into the following query, using the hashed index **230**:

Select 17

From 53

Where 47="0".

The converted query is processed in a traditional manner using the aliased database tables. For example, the installation program **137** will recognize from the query format that "53" is a table name, "17" and "47" are column names and "48" is a data element. A cursor of a row in table "53" will be created and the value "48" will be stored in column "47." Table "53" will then be searched, using the cursor, using a binary search method or any other well known search method. As can be seen from FIG. **2B**, the result set for this converted query comprises the data element "29." Referencing the index **230**, the string data element "Component_ Two" corresponding to the integer identifier "29" may be quickly located, i.e. by accessing index entry number **29**.

The uniformity of aliased database tables also allows for the direct addition of temporary rows and columns. FIG. **3** illustrates the concept of adding a temporary column **305** to a database table. As mentioned above, a relational database table is actually a two-dimensional array. In the situation where all fields in the array are uniform, it is a simple matter to expand the array to accommodate additional rows and column. In the exemplary embodiment, internal application program interfaces (APIs), such as "CreateColumn" and "CreateRow," are used to add temporary columns and rows to a database table. The internal APIs allow certain properties of the new column or row to be designated. One such property may be that the row or column is temporary. Externally, a temporary rows and columns are added using an SQL extension keyword, "TEMPORARY."

In the installation database context, temporary rows and columns may be useful when the result of a database query depends on dynamic external conditions. By way of example, prior to installation of a software product, it may be desirous to determine whether a given component of the software product is already installed on the target computer system **120**. The installation program module **137** may be programmed to scan the target computer system **120** to locate previously installed components. As shown in FIG. **3**, a temporary column **305** labeled "55" may be added to table "53" **215'** (the aliased Component_Table) to store a data element that indicates the installation status of each component of a software product **136**. If a previously installed component is located by the installation program module **137**, the integer identifier "56" is stored in temporary column "55" **305**. Note that entry number 55 in the index **230** is associated with the string "Status" and that entry number **56** is associated with the string "Installed."

While temporary rows and columns are in existence, data elements in both the temporary and permanent rows and columns may be directly accessed and modified. The data elements stored in those temporary rows and columns do not persist in the database. However, any modifications made to the data elements in the permanent rows and columns do persist, independent of the temporary rows and columns. Temporary data is also described in U.S. application Ser. No. 09/157,883, entitled "Management of Non-persistent Data in a Persistent Database" filed on Sep. 21, 1998 and owned by the assignee for the present application. This related application, U.S. application Ser. No. 09/157,883, is hereby hilly incorporated herein by reference.

The present invention further extends the functionality of an installation database system by providing methods for efficiently managing and manipulating installation database tables. For example, database tables may be merged using a "merge" or "union" operation of relational algebra. Software developers may thus create installation database tables in a modular fashion and merge the modules into a single installation database table when the software product **136** is complete. Also, installation databases may subsequently be created for patches or enhancements for a software product **136**. These subsequent installation databases may be shipped to an end-user and may be merged into the original installation database.

An exemplary merger of two aliased database tables is described in FIG. **4**. At step **402**, the first row of the first table is located. As previously mentioned, a row of a database table is identified by a key, which comprises one or more of the data elements in the columns of the row. At step **404**, a first cursor corresponding to the first table is populated with the data elements from the first row. Then, at step **406**, the contents of the first cursor are copied into a second cursor corresponding to the second table. At step **410**, the second table is searched to locate a row of data that is identified by the same key that is stored in the second cursor. If no matching key is found in the second table at step **412**, a new row is added to the second table and the data elements stored in the second cursor are copied into the new row at step **414**. However, if a matching key is located in the second database table, it is determined at step **416** whether the non-key data elements in the corresponding row are all identical to the non-key data elements in the second cursor. If the non-key data in the located row of the second table is not identical to the non-key data in the second cursor, a merge conflict is indicated at step **418** by flagging the conflicting row. Conflicting data elements are not merged into the second database table and may be called to the user's attention, so that such conflicts may be resolved. If the non-key data in the located row of the second table is identical to the non-key data in the second cursor, no action is taken with respect to that data.

At step **420**, a determination is made as to whether a next row exists in the first table. If a next row does exist in the first table, the next row is located at step **422**, the data stored therein is copied to the first cursor at step **404** and the above-described method is repeated. If no next row exists in the first table, the method is terminated at step **424**. To support merging and other database extensions, additional table and cursor operations have been implemented in addition to the standard insert, update, and delete. The merge operation performs the logic above for merging data into an existing table. Additional operations are: assign, which forces in data whether or not the key exists; replace, which allow primary keys to be edited; and several row validation operations. Externally these are invoked via the MsiView-Modify API, using specific enumeration values, along with the data record.

Installation database systems may also benefit from the "transform," or "difference," operation of relational algebra. This operation determines the differences between two databases and creates a temporary file of only those differences. A database transform is a recording of the differences

between two database tables. A database transform may be stored using a stream of changed data for each non-identical table.

Database transforms become very useful in providing customized installations, in synchronizing a program module to operate in different languages and in providing "patches" to fix various bugs in a program module. For example, an installation database may be modified to effect a custom installation of a software product **136**. A database transform may be created to record the differences between the original installation database tables and the modified database tables. Subsequently, the modifications recorded in the database transform may be applied to another copy of the original installation database, so that another end-user may obtain the same custom installation without having to manually perform the customization.

As another example, an English version and a Spanish version of a software product **136** may be created, each having its own installation database. A database transform may be created to record the differences between the English version and the Spanish version. Subsequently, the database transform may be applied to an installation database table corresponding to an updated English version of the software. The resulting "transformed" database will produce an updated Spanish version of the software product **136**.

FIG. **5** describes the creation of an exemplary database transform for recording the differences between a reference database table and a current database table. At step **502**, the first row of a current table is located. At step **504**, a current cursor corresponding to the current table is populated with the data elements from the first row. Then, at step **506**, the contents of the current cursor are copied into a reference cursor corresponding to a reference table. At step **510**, the reference table is searched to locate a row of data that is identified by the same key that is stored in the reference cursor. If no matching key is found in the reference table at step **512**, a row is added to a database transform and the data elements stored in the reference cursor are copied into the database transform at step **514**. However, if a matching key is located in the reference database table, at step **515** the corresponding row in the reference table is flagged to indicate that it has been processed. Then, at step **516** the data elements in the corresponding row that are not identical to the data elements in the reference cursor are copied into the transform at step **514**. If the data in the located row of the reference table is identical to the data in the reference cursor, no action is taken with respect to that data. At step **520**, a determination is made as to whether a next row exists in the current table. If a next row does exist in the current table, the next row is located at step **522**, the data stored therein is copied to the current cursor at step **504** and the above-described method is repeated. If no next row exists in the current table, the method proceeds to step **523**.

At step **523**, the primary keys of any non-flagged rows from the reference table are copied into the transform and are indicated to be rows that have been deleted from the current database table. Thus, when the database transform is applied to a new database table, the non-flagged rows will be deleted therefrom. In a similar manner, columns that are not present in the reference database may be registered in the transform as columns to be created at the time the transform is applied. Likewise, tables that are added or removed are registered in the transform, such as is necessary to be able to recreate a new database by applying the transform to the reference database. Following completion of step **523**, the method ends at step **524**.

Those skilled in the art will appreciate that certain principles of the present invention, while particularly beneficial to aliased installation databases, may be applied to any relational-type installation databases. For example, the concept of directly adding temporary rows and columns to a relational database table does not require that the data elements in the database table be aliased with integer identifiers. Furthermore, the concept of directly adding temporary rows and columns to a relational database table is not limited to installation database. Similarly, the concepts of providing upgrades, patches, enhancements, etc. for a software product by manipulating installation database tables through operations such as merges and transforms does not require the installation database tables to be aliased.

In view of the foregoing, it will be appreciated that the present invention provides optimizations for the process of installing a software product onto a target computer system. Still, it should be understood that the foregoing relates only to the exemplary embodiments of the present invention, and that numerous changes may be made thereto without departing from the spirit and scope of the invention as defined by the following claims.

I claim:

1. A method for reducing the persistent size and optimizing the performance of a relational database system, comprising the computer implemented steps of:

storing in an index a copy of each non-integer unique data element that would occur in a database table;

sequentially assigning to each entry in the index a corresponding integer identifier;

populating the database table with the corresponding integer identifier for each non-integer unique data element so that the database table uniformly contains integers; and

keeping a reference count for each non-integer unique data element to control the data element's lifetime.

2. The method of claim **1**, further comprising the steps of:

accepting a database query containing a queried data element;

searching the index to determine the integer identifier corresponding to the queried data element;

in response to determining the integer identifier corresponding to the queried data element, searching the database table to locate a record containing the integer identifier corresponding to the queried data element; and

processing the query.

3. The method of claim **2**, wherein the index is a hashed index; and

wherein searching the index comprises evaluating a hashing function.

4. The method of claim **2**, wherein the step of searching the database table comprises:

creating a cursor for the database table;

populating the cursor with information from the database query and the integer identifier corresponding to the queried data element; and

using the populated cursor to perform a binary search on the database table, so as to identify the corresponding integer data element in the database table.

5. The method of claim **1**, further comprising the step of adding a temporary row to the database table to store integer identifiers corresponding to temporary data elements.

6. The method of claim **1**, further comprising the step of adding a temporary column to the database table to store integer identifiers corresponding to temporary data elements.

7. The method of claim 5, further comprising the step of:

modifying the integer identifiers corresponding to data elements and temporary data elements;

wherein the modifications to integer identifiers corresponding to the data elements persist, but the modifications to the integer identifiers corresponding to temporary data elements do not persist.

8. A computer-readable medium having stored thereon the computer-implemented steps of claim 1.

9. A method for reducing the persistent size and optimizing the performance of a relational database system, comprising the computer implemented steps of:

storing in an index a copy of each non-integer unique data element that would occur in a database table;

sequentially assigning to each entry in the index an integer identifier;

populating the database table with the integer identifiers corresponding to each data element;

keeping a reference count for each data element to control its lifetime; and

adding a temporary column to the database table to store a plurality of temporary data elements, the temporary data elements not persisting in the database table.

10. The method of claim 9, further comprising the step of adding a temporary row to the database table to store the temporary data elements.

11. A method for reducing the persistent size and optimizing the performance of a relational database system, comprising the computer implemented steps of:

storing in an index a copy of each non-integer unique data element that would occur in a database table;

sequentially assigning to each entry in the index an integer identifier;

populating the database table with the integer identifiers corresponding to each data element;

keeping the reference count for each data element to control the data element's lifetime;

adding either a temporary row or a temporary column to the database table to store integer identifiers corresponding to temporary data elements; and

modifying the integer identifiers corresponding to data elements and temporary data elements; wherein the modifications to integer identifiers corresponding to the data elements persist, but the modifications to the integer identifiers corresponding to temporary data elements do not persist.

12. The method of claim 11, further Comprising the steps of:

accepting a database query containing a queried data element;

searching the index to determine the integer identifier corresponding to the queried data element;

in response to determining the integer identifier corresponding to the queried data element, searching the database table to locate a record containing the integer identifier corresponding to the queried data element; and

processing the query.

13. The method of claim 12, wherein the index is a hashed index, and wherein searching the index comprises evaluating a hashing function.

14. The method of claim 12, wherein the step of searching the database table comprises;

creating a cursor for the database table;

populating the cursor with information from the database query and the integer identifier corresponding to the queried data element; and

using the populated cursor to perform a binary search on the database table, so as to identify the corresponding integer data element in the database table.

15. A computer program product comprising a computer readable medium carrying computer executable instructions that implement a method for reducing the persistent size and optimizing the performance of a relational database system, the method comprising:

storing in an index a copy of each non-integer unique data element that would occur in a database table;

sequentially assigning to each entry in the index an integer identifier; populating the database table with the integer identifiers corresponding to each data element;

keeping the reference count for each data element to control the data element's lifetime;

adding either a temporary row or a temporary column to the database table to store integer identifiers corresponding to temporary data elements; and

modifying the integer identifiers corresponding to data elements and temporary data elements; wherein the modifications to integer identifiers corresponding to the data elements persist, but the modifications to the integer identifiers corresponding to temporary data elements do not persist.

16. The computer program product of claim 15, the method further comprising:

accepting a database query containing a queried data element;

searching the index to determine the integer identifier corresponding to the queried data element;

in response to determining the integer identifier corresponding to the queried data element, searching the database table to locate a record containing the integer identifier corresponding to the queried data element; and

processing the query.

17. The computer program product of claim 16, wherein the index is a hashed index, and wherein searching the index comprises evaluating a hashing function.

18. The computer program product of claim 16, wherein searching the database table comprises:

creating a cursor for the database table;

populating the cursor with information from the database query and the integer identifier corresponding to the queried data element; and

using the populated cursor to perform a binary search on the database table, so as to identify the corresponding integer data element in the database table.

* * * * *

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 6,606,618 B2
DATED : August 12, 2003
INVENTOR(S) : John C. Delo

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Title page,
Item [57], **ABSTRACT**,
Line 9, after "performance" please insert -- . --

Column 1,
Line 36, after "databases" please delete "."

Column 2,
Line 41, after "columns to" please delete "a"

Column 7,
Line 2, after "that" please delete "arc" and insert -- are --
Line 2, after "as a unit" please delete "," and insert -- . --
Line 4, after "to the user" please delete "," and insert -- . --
Line 13, after "present application" please delete "," and insert -- . --
Line 57, after "up to" please delete "64 k" and insert -- 64k --

Column 8,
Line 59, before "may be" please delete "in"

Column 9,
Line 4, after "seeks the" please remove "$Component_{13}Name$" and insert
-- Component_Name --
Line 34, before "In the exemplary" please delete "column." and insert -- columns. --
Line 40, after "Externally," please delete "a"

Column 10,
Line 6, before "incorporated" please delete "hilly" and insert -- fully --

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

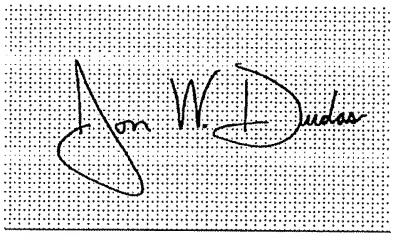
PATENT NO. : 6,606,618 B2
DATED : August 12, 2003
INVENTOR(S) : John C. Delo

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 13,
Line 47, after "further" please delete "Comprising" and insert -- comprising --

Signed and Sealed this

Twenty-fifth Day of May, 2004



JON W. DUDAS
Acting Director of the United States Patent and Trademark Office